Welfare Changes in Developing Economies: Technology Adoption, Cost-of-living, and Nutrition Improvement

by

Soye Shin

(Under the Direction of Nicholas Magnan and Chen Zhen)

Abstract

This dissertation comprises three essays on welfare changes among individuals in three different developing countries: Kenya, China, and Tanzania. Applying Prospect theory to experimental data, the first essay tackles a reason of the low uptake rate of weather index insurance among smallholder farmers although this risk management tool can prevent significant income losses due to rainfall shocks. The second essay highlights the importance of product variety and quality in constructing a cost-of-living index, a commonly used metric to assess consumers' welfare. The third essay investigates how households' nutrient intake changes in response to variations in food prices and total expenditures by estimating food demand systems and identify important foods for improved diet quality.

Index words: Weather Index Insurance, Prospect Theory, Cost-of-Living Index, Quality and Variety Bias, Food Demand System, Nutrition Intake

Welfare Changes in Developing Economies:

Technology Adoption, Cost-of-living, and Nutrition Improvement

by

Soye Shin

B.A., Korea University, Korea, 2010

M.S., Korea University, Korea, 2012

A Dissertation Submitted to the Graduate Faculty of The University of Georgia

in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2019

Welfare Changes in Developing Economies:

Technology Adoption, Cost-of-living, and Nutrition Improvement

by

Soye Shin

Major Professors: Nicholas Magnan
Chen Zhen

Committee: Ellen McCullough
Gregory Colson

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2019

Dedication

To my wonderful parents Myung-hee Kim and Dong-keun Shin. Their unconditional love and support throughout my intellectual quests keep me strong and courageous in every challenging moment. I owe to them everything I am, everything I have become, and everything I have accomplished.

To my Lord. I am who I am all by His Grace.

God chose things despised by the world,
things counted as nothing at all,
and used them to bring to nothing
what the world considers important.
As a result, no one can ever boast in the presence of God.

— (1 Corinthians 1:28-29)

Acknowledgments

I would like to express my deep gratitude to my two committee chairs, Dr. Nicholas Magnan and Dr. Chen Zhen. Their guidance and direction play an essential role in sharpening my critical thinking as well as writing and developing my early career as an independent researcher. I am also indebted to Nick for the field experiment opportunity in Kenya and to Chen for the provision of the Chinese scanner data which is only commercially available.

I also acknowledge Dr. Ellen McCullough for encouraging me to expand my research interests to health-related issues in developing countries through the Harnessing Food Demand Systems project. In particular, Ellen provided invaluable feedback and gave me the freedom to be creative, share ideas, and feel ownership of our work along the way. I also thank Dr. Gregory Colson for his service on my committee and advice on my professional development.

The first essay using the experimental data in Kenya was funded by 3ie through Phase 1 of Thematic Window 13: Agricultural Insurance: Assessing the Impact of Programmes Targeted at Smallholder Farmers project. I acknowledge Dr. Sarah Janzen and Dr. Conner Mullally for their critical feedback on the first essay. I would like to thank our implementing partners for supporting the research, particularly the government of Kenya's KCEP-CRAL leadership team, World Agroforestry (ICRAF), Acre Africa and the Kenya Promotions and Marketing Company (KPMC). I wish to acknowledge the contributions of Karl Hughes, Judith Odoul, Moses Abukari, Joseph Alulu, Dennis Kaburu, Dickson Osewe and Patrick Sampao for assistance with implementation at various stages of the research project. Finally, I am grateful to the Bill & Melinda Gates Foundation and the UK Department for International Development for funding the last essay through the Harnessing Food Demand Systems for Improved Nutrition in Sub-Saharan Africa project.

Table of Contents

List of Tables

List of Figures

Chapter 1

Demand for weather index insurance among smallholder farmers under prospect theory

1.1 introduction

Smallholder farmers are extremely vulnerable to climate variability which hampers crop and livestock productivity (Alinovi et al., 2010). For instance, significant income losses due to rainfall shocks result in sharp reductions in consumption (Dercon, 2004; Dercon and Christiaensen, 2011) that could cause adverse long-term effects on health and economic welfare. In developing countries, however, standard indemnity-based insurance is rarely available for smallholder farmers. As a result, weather index insurance (WII) has drawn much attention from researchers, development practitioners, and policymaker, as a promising way for smallholders to insulate themselves from climatic risks. Despite its potential merits, uptake of WII has been low (Giné and Yang, 2009; Binswanger-Mkhize, 2012; Cole et al., 2013; Hill et al., 2016; J-PAL, 2016), leaving economists to investigate why.

This paper attempts to provide a reason of the low uptake rate by exploring the relationship between risk preferences and insurance demand. Among others,[1]a handful of studies find that demand for WII appears to decrease with risk aversion (Giné et al., 2008; Hill et al., 2013; Cole et al., 2013). This finding is surprising in that Expected Utility theory (hereafter EU), which these studies have relied on, cannot predict an inverse relationship between insurance uptake and risk attitudes if insurance is perceived as a risk reducing tool. One possibility to explain this puzzling relationship is that farmers may view purchasing standalone WII as adopting new technology rather than a risk hedging tool (Hill et al.,

[1] See Carter et al. (2014) for review.

2013). This idea is based on empirical evidence that the main feature of index insurance, basis risk —the mismatch between actual losses and insurance payouts —is a relatively new and complex concept (Gaurav et al., 2011; Giné et al., 2013). However, this relationship may be a result of examining insurance purchasing decisions with a single risk parameter (i.e., risk aversion), for which alternative approaches are needed (Babcock, 2015).

Responding to this call, we use Prospect Theory (Tversky and Kahneman 1979, 1992; hereafter PT) to characterize individual risk preference. PT has been viewed as the most widely-used alternative to EU by shaping the utility function with three risk preference parameters: risk aversion, loss aversion, and non-linear probability weighting, which allows us to provide a richer view of the relationship between risk preferences and insurance demand. The loss aversion parameter, the main focus of this paper, captures an individual's degree of sensitivity to losses compared to gains of the same magnitude. What constitutes losses and gains depends not on absolute levels of wealth but a reference point. The reference point provides a "status quo" relative to which individuals evaluate gains and losses in their decision making. Individuals then weigh losses more heavily than gains in maximizing expected utility rather than simply maximizing the expected utility of final wealth. Choice of the reference point and the degree of loss aversion, therefore, are key factors to explain decision-making under uncertainty, including demand for index insurance.

We build a PT-based theoretical model and introduce two reference points to obtain theoretical predictions of WII demand. The difference between the two reference points is whether individuals view an insurance premium as a sunk-cost to reduce future risk or a loss. Guided by the model, we explore the empirical relationship between insurance demand and three dimensions of PT risk preferences. We then examine which theoretical prediction is consistent with empirical evidence that we find.

For empirical analysis, we collect experimentally-derived PT-based parameters and WII demand from 239 smallholder farmers in Kenya's arid and semi-arid lands (ASALs) in partnership with a Kenyan insurance provider widely working in East Africa. We first randomly assign one of two WII products (one with lower basis risk than the other) to the farmers and provide a brief explanation of the corresponding assigned product.[2] We then elicit parameters for risk aversion, loss aversion, and non-linear probability weighting using the methodology developed by Tanaka et al. (2010; hereafter TCN). Finally, we elicit demand at seven different prices using a multiple price list auction Andersen et al. (2006), which is a modified version of a Becker-DeGroot-Marschack (BDM) auction (Becker et al., 1964). The rich demand data allows us to examine if farmers' risk parameters respond to demand differently at various prices. One caveat is that very few farmers who purchase insurance through the auction actually purchased during our follow-up visits, rendering the data no more credible than that from a hypothetical auction.

We find that a PT framework performs better than an EU framework in explaining index insurance demand. Our empirical evidence shows that farmers are more likely to purchase insurance when they are more risk averse and less loss averse. The negative effect of loss aversion on insurance demand is consistent with our theoretical model irrespective of our choice of reference points. This effect potentially explains the inverse relationship between index insurance uptake and risk aversion that previous studies report. To date, these studies elicit risk aversion using a Binswanger-type game (Binswanger, 1980) where one can possibly confound risk aversion and loss aversion. In this regard, our finding suggests that the perplexing inverse relations between EU risk aversion and index insurance demand could be from a negative effect of loss aversion.

This paper also provides experimental evidence on the importance of a pay-upfront insurance premium by evaluating the negative effect of loss aversion at different insurance premiums. Examining which reference point explains the farmers' WII demand better, we find that the negative effect of loss aversion becomes more pronounced when farmers are offered insurance at high premium levels. Our results indicate that smallholder farmers consider an

[2] Because it is inappropriate to sell a lower quality product in the presence of a higher quality product, farmers purchasing the lower quality product with higher basis risk through the auction were given a free upgrade to the better product after the experiment.

upfront payment of premium as a loss as opposed to a sunk-cost to reduce future uncertainty by ruling the insurance premium out of their reference points.

We can interpret this finding in two ways. First, the sample farmers who have limited experience with insurance may have incorrect/narrow views about insurance where an upfront payment of insurance premium is not considered as a purchase of "risk-hedging device (e.g., cost to reduce consumption risk at harvest due to a drought)" but as an investment that is expected to give some returns to them, irrespective of a rainfall shock. Another possibility suggested by Casaburi and Willis (2018) is that smallholder farmers are highly likely to be liquidity constrained with cyclical incomes and therefore paying the premium upfront is not a marginal expenditure for them. This upfront payment could reduce liquidity for self-insurance *before* harvest, which may appear more important than risk management *at harvest* for farmers *at the moment of payment*. If this is true, farmers are likely to feel a loss when they pay premiums upfront. This is especially the case for WII products where there is a time gap between cost and benefit of the insurance products *and* the benefit is not even always guaranteed when rainfall shocks are realized due to basis risk.

We cannot argue which explanation is more compelling. Farmers may have a narrow frame of WII in which they do not account for the effects of insurance on overall income or consumption in their decisions (Brown et al., 2008) but focus on the outcome from the insurance contract itself. Conversely, they may consider expected wealth after the outcome of the insured event by viewing WII as a risk-management tool rather than a simple investment or lottery (Babcock, 2015). Our theoretical model and data do not allow us to test how farmers frame WII. In both cases, insurance premiums affect a marginal effect of loss aversion on insurance demand when the reference point excludes the premiums. Also, we elicit WII demand from farmers during a harvest season when they have the least liquidity constraints. The timing of the experiment could diminish the negative effect of liquidity constraints on insurance take-up while it is certainly possible that the poor are still susceptible to liquidity constraints even at the time of harvest. To summarize, while it is ambiguous which case explains heterogeneous effects of loss aversion by insurance premiums better, both consistently suggest that pay-upfront insurance premiums matter in smallholder farmers' insurance purchase decisions.

This paper adds to several strands of literature. First, an extensive literature has shown that PT performs better than EU in explaining individuals' seemingly suboptimal insurance choices[3] including low deductibles from auto/home insurance and low coverage levels of agricultural insurance (Kőszegi and Rabin 2006, 2007, 2009; Sydnor 2010; Bocquého et al., 2013; Babcock 2015). These studies focus on decisions of indemnity-based insurance. We show that index insurance demand is also better captured under a PT framework as opposed to an EU framework by closing a knowledge gap around the relationship between index insurance demand and a single EU risk parameter. We show that the previous findings are not counterintuitive but a result of a confounding effect.

As Barberis (2013) notes, how to incorporate reference points remains a key aspect to develop applications of PT in economics because it is often challenging to determine a reference point. Most PT related studies consider only one reference point, which substantially limits knowledge of how people subjectively process risky outcomes. A few recent exceptions use multiple possible reference points to obtain predictions that are close to observed choices (Babcock, 2015; Tonsor, 2018). This study attempts to fill this knowledge gap by assessing different reference points. We derive the predictions of PT under two different reference points (and therefore two different definitions of the domains of gains and losses) and test these predictions using experimental data.

Finally, several studies show that an insurance premium plays an important role in insurance demand with different angles. Focusing on a discontinuous preference for certainty, Serfilippi et al. (2016) design an unconventional index insurance where the payment of the premium is waived in bad years, showing that average willingness to pay for this product is 10% higher than a conventional index insurance with the same net insurance payment

[3]See, among other, Johnson et al. (1993); Grace et al. (2003); Barseghyan et al. (2013); Du et al. (2016)

in good and bad states, respectively. Casaburi and Willis (2018) provides experimental evidence that charging the premium at harvest leads a large increase in crop insurance take-up, implying that the timing of payment is an important determinant of insurance demand. Liu et al. (2017) also show similar results for hog insurance in rural China. Taking a different approach, we show that smallholder farmers feel a loss when they pay the premium upfront, which dampens index insurance demand. Our finding is contrast to the "no loss in buying" hypothesis proposed by Novemsky and Kahneman (2005),[4]suggesting that choices under uncertainty should be carefully analyzed in different contexts.

The rest of this paper proceeds as follows. Section 1.2 introduces the theoretical model. Section 1.3 describes the sample and weather index insurance products. Section 1.4 offers the experimental design and data collection process, including the follow-up visits to complete insurance sales. Section 1.5 presents results and Section 1.6 offers further discussion and some concluding remarks.

1.2 Theoretical model

As previously mentioned, a long literature shows anomalies of insurance purchasing behavior that are inconsistent with EU (e.g., the preference for low deductibles or sub-optimal coverage) can be explained by preferences in alignment with PT. Under the PT framework, an individual's utility function consists of not only risk aversion (as does EU) but also loss aversion and non-linear probability weighting, which is specified in (1.1) following Kahneman and Tversky (1979, 1992).

[4]They argue that money given up as part of a transaction is a "cost" and does not induce the psychological pain of a "loss". Their argument is based on an experimental data from US college students and paid-participants with inexpensive goods (e.g., mugs). Similarly, Kőszegi and Rabin (2006, 2007) segregate feelings about money given up for a purchase from attitudes toward unexpected losses in their framework.

$$
\begin{aligned}
U &= \sum_{i=1}^{n} \nu(x_i)\pi(p_i) \\
\text{where} \quad \nu(x_i) &= \begin{cases} x_i^{(1-\sigma)} & \text{for } x_i > 0 \\ -\lambda(-x_i)^{(1-\sigma)} & \text{for } x_i < 0 \end{cases} \quad \text{and} \quad \pi(p_i) = \exp\left(-(-\log p_i)^{\alpha}\right)
\end{aligned} \tag{1.1}
$$

In (1.1), there are n outcomes with their corresponding values, x_i that can be positive or negative. p_i denotes an associated probability of each x_i that is translated into a subjective probability according to a decision weight function $\pi(\cdot)$ derived by Prelec (1998). $\nu(\cdot)$ is the value function that is contingent upon whether x_i is positive or negative. The parameter σ describes the degree of risk aversion, with $\sigma > 0$ for a risk averse individual, $\sigma = 0$ for a risk neutral individual, and $\sigma < 0$ for a risk taking individual. The parameter λ defines the degree of loss aversion, with $\lambda > 1$ for a loss averse individual. The parameter α represents the non-linear probability weighting measure, with $\alpha < 1$ for an individual overweighting a small probability of unlikely extreme outcomes. Note that the PT model reduces to EU when $\alpha = \lambda = 1$. We will test the validity of the PT approach as opposed to the EU approach in 1.5.1.

As discussed in the Introduction, a reference point plays an important role to decide if an argument of the value function lies on the positive or negative domain. Accordingly, x_i indicates changes in nominal outcome values from a reference point. If a nominal value of an outcome denoted by m_i is greater (less) than a chosen reference point, R_i, an individual feels a gain (a loss). Accordingly, the value function and decision weighting function are defined as

$$
\nu(x_i) = \begin{cases} (m_i - R_i)^{(1-\sigma)} & \text{for } m_i > R_i \\ 0 & \text{for } m_i = R_i \\ -\lambda(R_i - m_i)^{(1-\sigma)} & \text{for } m_i < R_i \end{cases} \quad \text{and} \quad \pi(p_i) = \exp\left(-(-\log p_i)^{\alpha}\right) \tag{1.2}
$$

In (1.2), we can see how a reference point is incorporated into the loss aversion parameter. As the reference point increases, the probability of a particular outcome being a loss increases.

Since the loss aversion parameter (λ) causes higher values to be assigned to a loss in absolute magnitude than on the same amount of a gain, the pain of losses is likely more and more significant as the reference point increases with other things being constant.[5]

Inspired by Serfilippi et al. (2016) and Casaburi and Willis (2018), we choose two reference points the difference of which is the paid insurance premium. We denote W as the initial wealth or the initial wealth plus the expected revenue from farming, depending on how farmers frame insurance[6] and ρ as an insurance premium. The first reference point is W and the second reference point is $W - \rho$. If farmers view the insurance premium as a loss then W will be their reference point while $W - \rho$ will be their reference point when they perceive the premium as a sunk-cost in exchange for reduced future uncertainty.

Given a choice of the reference point, we describe farmers' WII purchase choice within a PT framework by specifying simplified cases that fall into either the domain of gains or losses. We denote L as a crop loss, γL as an insurance payout with $\gamma \in (0, 1]$ such that $\gamma L > \rho$.[7] We define p as a probability of rainfall shocks, q_1 as a probability of basis risk conditional on a rainfall shock realization (i.e., negative basis risk), and q_2 as a probability of basis risk conditional on the absence of a rainfall shock (i.e., positive basis risk). We start with the first reference point, W.

Four cases exist when a farmer purchases insurance with a reference point W.

[5] Due to the convexity of the value function in the region of losses, the sensitivity of a loss is diminishing as the reference point increases. However, we think this effect is marginal on insurance choices because as Barberis (2013) mentions, an individual facing a loss that represents a large fraction of wealth will be very sensitive to any additional losses, which is highly likely to be the case for our sample farmers. The convexity does not capture this psychological intuition.

[6] With a broad frame of WII, farmers consider expected wealth after the rainfall outcome is realized, and therefore the expected profit from farming is relevant. In contrast, farmers who have a narrow frame of WII evaluate the value of insurance on the basis of insurance gains or losses, hence initial wealth is only considered.

[7] The main findings of the paper do not change when we assume that an insurance payout equals to L. The WII products that we sold provide an insurance payment to cover full losses that a farmer has, with a significantly small probability even when the farmer purchases full insurance and no basis risk is realized. For the case of $\gamma L < \rho$, we will see a positive effect of a marginal change in risk aversion on insurance demand when the reference point is W. Refer to Appendix A.

1. There is a rainfall shock and insurance payout γL is realized (i.e., no negative basis risk). The probability of the state is $p(1-q_1)$. In this case, the final wealth of the farmer is $W-\rho-L+\gamma L$ and the farmer feels a loss of as much as $\rho+L(1-\gamma)$ given his reference point.

2. There is a rainfall shock but no insurance payout is realized (i.e., negative basis risk). The probability of the state is pq_1. In this case, the final wealth of the farmer is $W-\rho-L$ and the farmer feels a loss equal to $\rho+L$.

3. There is no rainfall shock and no insurance payout is realized (i.e., no positive basis risk). The probability of the state is $(1-p)(1-q_2)$. In this case, the final wealth of the farmer is $W-\rho$ and the size of the loss that the farmer feels is ρ.

4. There is no rainfall shock and insurance payout γL is realized (i.e., positive basis risk). The probability of the state is $(1-p)q_2$. The final wealth of the farmer is $W-\rho+\gamma L$ in which case the farmer feels a gain equal to $(\gamma L-\rho)$ (assuming that γL is greater than ρ).

In summary, there are four possible outcome values that x_i takes with corresponding probabilities.

$$x_i = \begin{cases} -\rho - L(1-\gamma) & \text{with } p(1-q_1) \\ -\rho - L & \text{with } pq_1 \\ -\rho & \text{with } (1-p)(1-q_2) \\ -\rho + \gamma L & \text{with } (1-p)q_2 \end{cases} \tag{1.3}$$

Following (1.1) and (1.2), an individual's expected utility of an insurance purchase with the reference point of W is specified as follows.

$$\begin{aligned} U_1(\rho, L, \gamma; p, q_1, q_2) = & -\lambda(\rho + L(1-\gamma))^{(1-\sigma)}\pi(p - pq_1) - \lambda(\rho + L)^{(1-\sigma)}\pi(pq_1) \\ & - \lambda\rho^{(1-\sigma)}\pi(1 - p - q_2 + pq_2) + (\gamma L - \rho)^{(1-\sigma)}\pi(q_2 - pq_2) \end{aligned} \tag{1.4}$$

We standardize the utility of the status quo (i.e., no insurance) at zero. In this setting, a farmer will purchase insurance when the expected utility from the insurance product is $U > 0$ and will not when $U < 0$.

In the same way, we obtain an individual's expected utility when the chosen reference point is $W - \rho$.

$$U_2(\rho, L, \gamma; p, q_1, q_2) = -\lambda(L(1-\gamma))^{(1-\sigma)}\pi(p-pq_1) - \lambda L^{(1-\sigma)}\pi(pq_1) + (\gamma L)^{(1-\sigma)}\pi(q_2-pq_2) \quad (1.5)$$

Using (1.4) and (1.5), we can predict an individual's WII participation with respect to a marginal change in each risk preference parameter, which is shown in the following summary of comparative statistics. We find that the two models with different reference points provide different marginal effects in magnitude but the same qualitative predictions.

i $\frac{\partial U}{\partial \sigma} \lesseqgtr 0$: WII participation with respect to a change in risk aversion is ambiguous.

ii $\frac{\partial U}{\partial \lambda} < 0$: Loss averse farmers are less likely to take up a WII product.

iii $\frac{\partial U}{\partial \alpha} \lesseqgtr 0$: WII participation with respect to a change in nonlinear probability weighting is ambiguous.

Appendix A contains the derivations of the comparative statistics. Our theoretical predictions imply that if PT explains index insurance choices of smallholder farmers well, we will see that the loss aversion parameter dampens WII demand.

To examine how insurance choices differ by our choice of the reference point among two, we pay attention to the role of a paid insurance premium in the arguments of the value functions. Because the paid premium is the only difference between the two reference points, it can affect the extent to which a farmer feels a gain/loss, and therefore insurance purchase is worthwhile/not worthwhile under the first reference point W *but not* the second reference point $W - \rho$.

We can test it by examining how a marginal effect of loss aversion differs by insurance premiums under the two reference points, respectively. If an insurance premium matters, a negative effect of loss aversion becomes more pronounced when insurance is offered at high premiums. We can show this phenomenon by taking a second derivative of $\frac{\partial U}{\partial \lambda}$ with respect to insurance premium ρ for each reference point and compare signs of the marginal effect of loss aversion conditional on premiums across the two cases. Accordingly, we present the reference-specific formulas of ii in the summary of the comparative statistics and its second derivative with respect to ρ as follows.

$$
\begin{cases}
\frac{\partial U_1}{\partial \lambda} = -\pi(p - pq_1)(\rho + L(1-\gamma))^{(1-\sigma)} - \pi(pq_1)(\rho + L)^{(1-\sigma)} - \pi(1 - p - q_2 + pq_2)\rho^{(1-\sigma)} < 0 \\
\quad \frac{\partial^2 U_1}{\partial \lambda \partial \rho} = -(1-\sigma)[\pi(p - pq_1)(\rho + L(1-\gamma))^{\sigma} + \pi(pq_1)(\rho + L)^{\sigma} + \pi(1 - p - q_2 + pq_2)\rho^{\sigma}] \\
\qquad < 0 \qquad \text{when } \sigma < 1 \\
\qquad > 0 \qquad \text{when } \sigma > 1 \\
\frac{\partial U_2}{\partial \lambda} = -\pi(p - pq_1)(L(1-\gamma))^{(1-\sigma)} - \pi(pq_1)L^{(1-\sigma)} < 0 \\
\quad \frac{\partial^2 U_2}{\partial \lambda \partial \rho} = 0
\end{cases}
\tag{1.6}
$$

In the case of $R_i = W$, as long as the risk aversion parameter is less than 1, the magnitude of marginal disutility of loss aversion is larger when insurance premium increases. This result implies that an individual who is not extremely risk averse dislikes insurance products more when he/she faces a higher premium. In contrast, the result of the second derivative under $R_i = W - \rho$ is independent to changes in ρ, implying that the marginal effect of loss aversion is constant across different insurance premiums. Therefore, a heterogeneity analysis of loss aversion with the premium can provide insight about the reference point in insurance choices.

The ambiguous signs of the relationships between other two risk parameters (risk aversion and nonlinear probability weighting) and WII demand suggest that directions of the associations should be established empirically. We will examine how WII demand varies with risk aversion and nonlinear probability weighting and interpret results within a context of sample farmers.

1.3 Sample and Weather Index Insurance Products

The data used in this paper have been collected as part of the evaluation of a broader randomized controlled trial (RCT) carried out in collaboration with ACRE Africa (Agriculture and Climate Risk Enterprise, ACRE hereafter), an insurance provider widely working in East Africa. We target smallholder producers in Tharaka South sub-county of Tharaka Nithi who are vulnerable to climate shocks, especially droughts. Our study area is one of the Kenyan ASALs on which the government of Kenya has recently focused to enhance climate-resilient productivity with a large amount of policy funds called KCEP-CRAL. Most people in the study area are subsistence farmers growing sorghum or green gram, with a smaller number growing millet, maize, cow peas, or pigeon peas. In addition, they have no experience with WII before the study.

Prior to the study, an aggregator played a role as an intermediary between a large group of farmers and output markets by supplying a bundle of inputs on credit to smallholder farmers growing sorghum and/or green gram. To insulate herself from farmers' default risk due to weather shocks, the aggregator purchases WII for the value of inputs through ACRE, which farmers did not know. This insurance contract was developed by ACRE using ARC2 data in 2013. Because most farmers in our sample are working with the aggregator, we inform them that their input loans are already insured and provide them with the chance to buy insurance for the projected value of sorghum and/or green gram production *less* the value of insured inputs. We focus on sorghum insurance because sorghum is the main crop to lead a transition from subsistence to commercial farming, yet one of the most vulnerable crops to droughts.[8]

The insurance products used in this study were generated using Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data (Funk et al., 2015). As part of interventions for the broader RCT project, we design two WII policies that are differentiated only

[8]Therefore, we only provide green gram insurance to farmers who are not planning to plant sorghum in the upcoming season.

by basis risk. A high resolution (HR) product was generated using 5km x 5 km grid squares and a low resolution product (LR) was created using 10km x 10km grid squares. While the difference in demand for the LR and HR product is not the primary focus of this paper, we do randomly assign each farmer to one of the two products and account for this variation in our analysis when this variation is important in explaining the relationship between risk attitude and insurance demand. For ethical reasons, farmers purchasing the LR product through the auction were given a free upgrade to the HR product.

The insurance policy is complicated in that policy triggers and ways of calculating payouts with maximum amounts vary not only with crops but also growing stages (e.g., four stages from germination to pre-harvest at the pentad-level) and satellite locations (pixel).[9] ACRE shares with the research team their previous experience with farmers, saying that this is beyond the comprehension of farmers and often intimidates them. Through two pilot studies, we also learned that the target population has low levels of mathematical literacy. Therefore, we decide not to provide the study farmers with policy details.

We collect data including individual characteristics from 239 farmers, 191 of whom has previously worked with the aggregator. The others could begin working with the aggregator as part of the study. Table 1.1 presents the descriptive statistics for sample farmers we use for our analysis. Mean landholding among the sample population is 4.48 acres (with the median of 3.5 acres), and 88.7 percent of farmers experienced at least two droughts during the recent five agricultural seasons. The mean educational attainment is 7.2 years (with the median of 8 years) and 26 percent of farmers claimed to be illiterate. The age of participants ranges from 18 to 83 with the median of 38. Loan availability[10] for the next agricultural season and ownership of a formal saving account (regardless of active use) are used as proxies for liquidity constraints. For these variables, 17.6 percent of farmers said they would not be able to take loans and did not have a formal saving account.

[9] For more information on the policy design, see Appendix A.

[10] The loan availability excludes loans for inputs that the aggregator provides.

1.4 Experimental Design

We conducted the experiment and data collection in July 2017. The structure of each session was ordered as follows: (1) introduction to index insurance, (2) short survey on farmer and farm characteristics, (3) PT game, (4) experimental auction, (5) post experiment survey, and (6) game payout.

Three weeks later we conducted follow-up visits to complete transactions for those purchasing insurance through the auction. This considerable time gap was unavoidable given that all field-work had to cease leading up to and following the historically chaotic, and eventually nullified, 2017 Kenyan election. Most of the farmers did not fulfill their commitments they made in the auction, although the auction was meant to be binding. More discussions are in subsection 1.4.5.

As one session takes about three and a half hours, we conducted two sessions per day. In one session we offered the LR product and in the other we offered the HR product, determined at random. Under the sample frame that the aggregator provides, we initially had a plan to randomly assign individual farmers to each session, but this proved to be impossible. Therefore, lead farmers assigned individual farmers to a morning or afternoon session without knowing treatment types and any activities corresponding to the treatments. Each session took place in an open space with a group of approximately 16 farmers.

1.4.1 Information Session

We designed the information session to be similar to an ACRE's marketing strategy ("business as usual") as much as possible. The promoter, a person directly working with farmers for the aggregator, led the information session to minimize farmers' lack of trust in insurance products and following activities. He started the session with insurance needs within the farmers' contexts that this risk management tool helps farmers grow their business with output markets by increasing productivity of sorghum and green gram.

He also introduced the concept of basis risk by explaining that insurance payouts are determined not by the rainfall in individual farms but by the rainfall in a size of a square (either 5 x 5 or 10 x 10 km) containing those farms. To help farmers have a better sense of the size of the square, he took an example of a rectangular made by well-known local locations with approximately 5 km or 10 km being apart from each other. Farmers were explicitly told a negative basis risk event with an example but not a positive basis risk event. This is because we would like to remove a possibility that farmers might consider the insurance product as a lottery if they were explicitly informed that they could receive payouts even in the absence of crop losses. Again, due to the complexity of the insurance policy, information about probabilities and the triggering system was not provided.

1.4.2 Prospect Theory Game

As mentioned in the introduction, the PT game is a version of the TCN experiment. In the PT game, participants were given three series of pair-wise choices between two lotteries. The first two series contain 10 such choices, respectively, and the last series has 7 choices. Table 1.2 illustrates the game's entire payoff matrix. We reduced the number of game rounds for series 1 and 2 from 14 (the number of game rounds in TCN) to 10 for this game to save time.

Before the real game, we showed multiple progressive examples to help farmers better understand the game process. The game master first explained how the lottery works with a simple lottery with a single prospect by drawing a number between 1 and 10 from an envelope. For instance, if a lottery gives 10 Kenyan Shillings (KSH, 1 USD $\approx$ 100 KSH) when the drawn number is between 1 and 7, the game master drew a card and informed of which case farmers would receive 10 KSH. After practicing with another simple lottery with two prospects, the game master showed the following two lotteries and asked them which lottery they would choose.

Lottery A	Lottery B
10 if 1 2 3 4 5 6 7 20 if 8 9 10	20 if 1 2 3 4 5 6 7 8 9 10

If a farmer understands how the lottery works, he/she must know lottery B is less risky and always the better option. Once all farmers understood another example with two lotteries, farmers played 5 rounds of a practice series similar to the real game where lottery B is a riskier option and the risk degree becomes larger moving down the table. Farmers were told that each series would end once they switch to (or choose) lottery B, and the rationale of it (e.g., it is not rational that you switch back to a less risky option once you say you prefer a riskier lottery over a less risky).

The game master then demonstrated how payouts are determined. By randomly drawing a card, he first selected a series and then a round of that selected series. Another card between 1-10 was then drawn to determine the game payout from the selected round. The average total payout (184 KSH)[11] is small enough relative to the actual price of insurance to avoid income effects on insurance purchases.[12]

In each series, an enumerator asked the farmer round-by-round which lottery he/she preferred. Once the farmer chose lottery B, the enumerator recorded the switching point and proceeded to the next series. After the game, enumerators calculated total payouts and confirmed them with farmers but payments were made at the end of the whole experiment.

1.4.3 Auction

To elicit WII demand, we used a multiple price list auction (Andersen et al., 2006) in which farmers revealed their desired insurance coverage level at seven different prices. Once participants revealed how much insurance they would like to buy at each price, the sale price of insurance (premium) was randomly drawn from an envelope. When the drawn price was above the market price, we replaced it with the market price and adjusted the insurance

[11]Payouts were also calibrated to eliminate the situation where a farmer would pay money out of his/her pocket.

[12]At actuarially fair prices, insuring one acre of sorghum production costs approximately 3,600 KSH. We also check for income effects by regressing insurance demand on total payouts and find no effect.

quantity that farmers would buy accordingly. We did not tell farmers we would do this beforehand so that they would consider these higher prices to be possible.

The auction is semi-binding in the sense that farmers did not have to pay premiums right after the experiments, but they would not be allowed to purchase more than they agreed to in the auction during the follow-up visits. There are three reasons for this. First, it may not be reasonable to compel farmers to buy WII on the first day they learn the products. Second, it is not reasonable to expect farmers to have adequate cash on hand, or in their mobile M-Pesa accounts, to make a purchase on the day of the auction. In either case, forcing them to complete a transaction on the day of the auction would likely depress demand. Third, the experiment was conducted before the period over which agricultural insurance is normally purchased (i.e., at planting) due to constraints on grant spending and the need to complete fieldwork before the 2017 Kenyan general election.

We initially measured insurance quantities as total coverage amounts (in KSH) and premiums as a percentage of the amount covered. However, in pre-testing we found that the price stated in percentage terms is unfamiliar to farmers. We therefore stated premiums as the cost of a certain quantity of coverage. We offered units providing up to 5000 KSH of coverage at the following prices: 50, 150, 250, 350, 500, 750, and 1000 KSH.[13] This way, farmers only need to state how many units of coverage they would like to buy at each price (partial units were not an option unless their projected net production values are less than 5000 KSH). We confirmed that both enumerators and farmers had no difficulty understanding the process during a second pilot study. To help farmers better understand the auction process, we conducted a practice auction for cookies before the real auction.

The auction was conducted iteratively as follows. An enumerator first helped farmers calculate the projected net value of sorghum produced by subtracting insured input values (i.e., inputs provided by the aggregator) from the projected total value. Starting at the lowest price, the enumerator told a farmer the number of insurance units needed to insure

[13]These values are in percentages of 5000 KSH (1%,3%,5%,7%,10%,15%, and 20%). The market price is 15% of insurance coverage.

his/her total projected net value of produce and total cost of that many units, asking if the farmer would like to purchase those units of insurance. If the farmer says no, the enumerator decreases the number of units and the corresponding coverage level and again asks if the farmer would like to purchase the insurance. If he/she agrees to the purchase, the enumerator proceeds to the next price and does the same things until either all prices in the list have been covered or a price is reached at which the farmer does not wish to purchase any insurance.

After the actual insurance price was determined, the enumerator confirmed with the farmer that he/she had committed to purchase the amount of insurance at the selected price. Then, the enumerator left her/him with an informational pamphlet about index insurance with the price and quantity the farmer agreed to buy. The complete auction protocol can be found in Appendix A.

1.4.4 Post-survey

At the conclusion of the auction, we conducted a post-survey to collect farmers' attitudes towards WII and basis risk knowledge. The knowledge of basis risk was measured by asking four questions to farmers and counting their number of correct answers. The four questions with the format of True/False are as follows: 1. If I buy the insurance, I will always receive money back at the end of the season, regardless of the weather; 2. It is possible to receive an insurance payout even if I have received enough rain on my farms; 3. To determine how much rain has fallen, the insurer measures overall rainfall over larger squared areas, not at a single farm; 4. The smaller squared area the insurer uses to measure rainfall, the greater the similarity between rainfall measured for the squared area and rainfall on my farms.

The mean value of the number of correct answers of basis risk relevant questions is 2.97, and the median is 3. These are fairly high scores given that farmers were told about basis risk briefly in the information session. We attribute the high scores to the way enumerators elicit farmers' desired insurance coverage in the auction. At each price, enumerators told farmers that insurance quantities were the *maximum* payouts they could receive when the weather is bad. Although the concept of basis risk was not directly explained to farmers in this statement, we think that many farmers have learned the possibility of not fully recovering their insured losses with this insurance product.

Examining the portion of farmers who chose the right answer for each question, we find that the percentage of the farmers with correct answers (46%) is relatively much lower for the question related to positive basis risk (Question 2) than that of the other questions (from 76% to 93%). Therefore, we are inclined to believe that a positive basis risk event is unlikely to be heavily accounted for by sample farmers.

1.4.5 Follow-up visits

Prior to the follow-up visits in late August, all individuals committing to purchasing insurance were reminded via multiple text messages and phone calls of their commitment to purchase insurance via M-Pesa mobile money or in person. Because farmers were not paying, we organized two in-person visits and extended the payment deadline. In the end, only six of the 361 farmers (i.e., the overall number of farmers who said to purchase in the project) paid their premium. Even in light of evidence of low uptake rates of WII found in other studies, it is surprising that farmers did not purchase insurance at highly subsidized prices (e.g., the lowest premium equals to the 93% discount price of the market price of sorghum WII).

Most farmers cited the lack of cash after paying for pressing needs (e.g., school fees, medication) as the main reason for not being able to purchase. Indeed, payment was requested at the onset of the school year. In addition to the bad timing, we suspect that too much time passed (three weeks) between the intervention and planting time due to, as stated earlier, constraints on grant spending and fieldwork during the election.

Although WII demand from the auction did not translate to actual purchases, we can still learn about the relationship between risk preferences and demand under the assumption that the amount of hypothetical bias is equivalent to everyone irrespective of individuals' risk preferences. Because most sample farmers already built trust with the aggregator for the input bundles it is hard to believe that farmers did not take this experiment seriously at all therefore the bias is substantial. However, we cannot eliminate a possibility that the hypothetical bias systematically varies with risk preferences from the data. Therefore, any implications from our findings should take this possible limitation into account.

1.5 Empirical Results

1.5.1 Descriptive Analysis of Game Results

Using a combination of switching points from Series 1 and 2, we can estimate risk aversion (σ) and non-linear probability weighting (α) parameter. After obtaining an estimate of σ, we use the switching point in series 3 to identify the loss aversion parameter (λ) conditional on a value of σ. The distributions of risk preference parameters are presented in Figure 1.1. The distribution of non-linear probability weighting approximates to a normal distribution, but the distributions of other parameters do not.

The data shows that 64.4 percent of farmers are risk averse ($\sigma > 0$) and 66.5 percent exhibit loss aversion ($\lambda > 1$). Also, 69.9 percent of the sample overweights small probabilities and underweights large probabilities ($\alpha < 1$). The mean values of σ, α, and λ are 0.75, 0.89, and 3.99, respectively.[14] A test of the validity of EU (as opposed to PT) rejects the null hypothesis that $\alpha = 1$ and $\lambda = 1$ (i.e., the null hypothesis that the PT utility function collapse to EU) at the 0.1% level (p-value: 0.000). This result implies that the EU framework fails to adequately explain farmers' risk preferences, and therefore how risk preferences relate to insurance demand.

[14]These values are analogous to those found in other studies, implying the validity of the game. For the mean values of σ, α, and λ, TCN (2010) finds 0.59, 0.74, and 2.63 among Vietnamese farmers, Liu (2013) finds 0.48, 0.69, 3.47 among farmers in China, and Bocquého et al. (2013) finds 0.51, 0.65, and 3.76 among French farmers, respectively.

1.5.2 Effects of Risk Parameters on Insurance Demand

Basic Regression

We first test that WII demand decreases with the loss aversion parameter by regressing quantity demanded on risk parameters, controlling for insurance price and a list of covariates. We have seven observations, one at each price, for each farmer, for a total of 1,673. The model is specified as:

$$Q_{ij} = \beta_0 + \beta_1\sigma_i + \beta_2\alpha_i + \beta_3\lambda_i + \beta_4\rho_{ij} + \beta_5\rho_{ij}^2 + \beta_6 HR_i + \phi\mathbf{X_i} + \theta_k + \varepsilon_{ij}. \qquad (1.7)$$

In equation (1.7), Q_{ij} is the quantity of insurance demanded in KSH by farmer i at insurance premium level j. The three risk attitude parameters include σ_i (risk aversion), α_i (non-linear probability weighting), and λ_i (loss aversion). ρ_{ij} is the insurance premium at which demand was reported and we include the square of the insurance premium to capture diminishing marginal disutility of premium. HR_i is the indicator for treatment status of being offered the HR product and X_i represent subjects' characteristics including age, gender, household headship, years of education, a literacy indicator, land size, the number of droughts in the past five agricultural seasons, an indicator for loan availability, and indicator for having formal saving accounts,[15] acres of the insured crop, anticipated harvest values of the insured crop, and an indicator for the insured crop being sorghum. We also control for enumerator fixed effects (θ_k).

Effects of risk preference parameters from the estimation of equation (1.7) are in Table 1.3, column 1. We find that farmers are likely to buy more insurance when they are more risk

[15] A dummy for loan availability and having formal saving accounts (irrespective of actual use) are included to control for liquidity constraints and a potential relationship with insurance as an alternative risk-management tool (either as substitutes or complements) which certainly affect insurance purchases. One may argue that these variables are endogenous to risk preferences. However, we do not think this is a huge concern. Whether to have *ability* (not willingness) to take loans if needed is not dependent on risk attitudes. Also, almost all sample farmers are members of informal saving-groups through which they save, and a simple ownership of a formal saving account only indicates whether an individual has access to formal financial services, which would increase his/her financial liquidity.

averse and less loss averse. These effects are significant with a p-value of 0.034 and 0.050, respectively. A positive effect of risk aversion is consistent with the results that are observed in the indemnity-based insurance literature but inconsistent with outcomes from previous studies about index insurance. Inconsistency with the previous empirical evidence would not be surprising given that the previous evidence was found based on EU. As mentioned earlier, the inverse relation between WII demand and risk aversion is possibly due to the negative effect of loss aversion that is confounded with the EU risk aversion parameter. The negative effect of loss aversion among our sample farmers supports this confounding effect and, therefore suggests that PT does a better job than EU in explaining how risk attitudes influence index insurance demand.

Although the effect of the risk aversion parameter is ambiguous in the theory, we expect to see a positive effect on insurance demand among sample farmers. Guided by our model, we can show $\frac{\partial U}{\partial \sigma} > 0$ for both reference points when utility gain from the outcome due to positive basis risk is less than utility loss from the other outcomes in the loss domain. As explained, we only explicitly introduce a negative basis risk case (where farmers do not receive insurance payouts for their insured losses) in the information session. Combined with the relatively low percentage of the correct answer to the positive basis risk question, we think our sample farmers did not take positive basis risk into account considerably in insurance purchase decisions.

We also find from the positive effect of non-linear probability weighting (α) that individuals who overweight small probabilities are less likely to purchase index insurance than who do not. To show this result more easily, we create a dummy variable for those who place excessive decision weight on small probabilities (i.e., $\alpha < 1$) and run the same regression only by replacing the non-linear probability weighing measure with the dummy variable. The result is reported in column 2, showing the negative effect of the tendency to overestimate small probabilities on the index insurance demand. Then, the natural question that arises is what event the farmers perceive as an unlikely outcome.

In the literature on insurance and technology adoption, the unlikely outcome is normally a disaster or a negative shock against which individuals want to be insured or avoided. In our study, that is drought. However, we learn from the pre-survey that the likelihood of rainfall shocks that the sample farmers perceive is fairly large. Almost two-thirds of the sample farmers (64.44%) said they experienced more than three droughts in the past five agricultural seasons (60% chance of a drought). The proportion increases to 88.71% when we include farmers who experienced more than two droughts during this period (40% chance of a drought). Thus, among the sample farmers, the unlikely outcome is not a rainfall shock but could be a basis risk event, specifically the negative one.

Under this claim, the positive effect of α suggests that farmers who overweight the probability of a negative basis risk event would be those who are less likely to purchase index insurance. We can show in what case $\frac{\partial U}{\partial \alpha} > 0$ holds using our theoretical model. We only discuss the case when the reference point is $R_i = W$ as the same logic applies to the second reference point case ($R_i = W - \rho$).

$$\begin{aligned}\frac{\partial U_1}{\partial \alpha} =& -\frac{\partial \pi(p - pq_1)}{\partial \alpha}\lambda(\rho + L(1-\gamma))^{(1-\sigma)} - \frac{\partial \pi(pq_1)}{\partial \alpha}\lambda(\rho + L)^{(1-\sigma)} \\ & - \frac{\partial \pi((1-p)(1-q_2)}{\partial \alpha}\lambda\rho^{(1-\sigma)} \quad + \frac{\partial \pi(q_2 - pq_2)}{\partial \alpha}(\gamma L - \rho)^{(1-\sigma)}\end{aligned} \tag{1.8}$$

If farmers overweight the probability of a negative basis risk event when the probability of a drought is large (e.g., $p = 0.4$), all terms except the second term (the outcome of a negative basis risk event) are negative.[16] When the positive value of the second term is greater than the sum of the negative values from the other three terms, it must be true that $\frac{\partial U}{\partial \alpha} > 0$. It is also true that the larger the absolute values of $-(\rho + L)$ (the loss from a negative basis risk event) and $\frac{\partial \pi(pq_1)}{\partial \alpha}$ are, the more likely $\frac{\partial U}{\partial \alpha} > 0$ is to be the case.

Beyond the theoretical explanation of the effect of α, the above discussion illuminates a role of the probability of negative basis risk in the effects of risk parameters on demand. With the data available, we further examine it in the next subsection.

[16] For details, see the derivation in Appendix A.

Heterogeneity Analysis with Basis Risk Probabilities

Following the discussion in the previous subsection, it is worth exploring heterogeneous effects of risk parameters on insurance demand by basis risk probabilities using the variation in insurance resolutions (LR or HR). Combining the theoretical model and the study context, we build prior directions of these effects as hypotheses to test. Continuing the discussion of the case of $\frac{\partial U}{\partial \alpha} > 0$, we first demonstrate the theoretical prediction of the effect of non-linear probability weighting conditional on HR.

In the previous subsection, we know that a rise in the absolute value of $\frac{\partial \pi(pq_1)}{\partial \alpha}$ increases the probability of having $\frac{\partial U}{\partial \alpha} > 0$ given that the probability of a rainfall shock that the sample farmers have experienced is large and the probability of negative basis risk (q_1) is relatively small. In this setting, the value of $\frac{\partial \pi(pq_1)}{\partial \alpha}$ becomes larger in absolute sense with decreasing q_1 when other parameters are being constant. This relation suggests that the positive effect of the non-linear probability weighting measure is greater when farmers are introduced HR than LR.

The decrease in the probability of negative basis risk also plays a role in the effect of loss aversion on insurance demand. Intuitively, we expect that loss averse farmers are likely to buy more insurance when they have a chance to buy HR as opposed to LR. This intuition can be proven by the theoretical model as well. Consider the partial derivative formulas in equation (1.6). Suppose the probability of a negative basis risk event decreases (i.e., upgrade insurance from LR to HR) while every parameter is constant.[17] The value of the decision weight function ($\pi(\cdot)$) decreases for the negative basis risk event whereas the values increase for the other non-basis risk event(s). Unless the probabilities of the negative basis risk event for HR and LR are substantially large (e.g., 0.7 and 0.8), the amount of decrease always exceeds the amount of increase, therefore, the negative effect of loss aversion becomes smaller.

[17]The probability of positive basis risk also decreases when the insurance product is upgraded to HR. However, the decrease in q_2 does not affect the marginal effect of loss aversion on insurance demand as the positive basis risk event does not involve a loss.

For the same reason related to value changes of the decision weight function, the theoretical model indicates that risk averse farmers with HR are likely to buy less insurance than those with LR. With the product upgrade, the values of $\pi(\cdot)$ for the positive and negative basis risk events decrease and the decreased value is greater than the increased value from the non-basis risk events. Assuming the positive basis risk event is the least relevant to farmers' insurance decision making, the positive marginal effect of risk aversion, which is empirically shown among the sample farmers, decreases when farmers are provided HR as opposed to LR.

Based on the discussion above, we can examine whether the effects of each risk parameter vary by HR by adding interaction terms between the risk parameters and HR to equation (1.7).

$$\begin{aligned} Q_{ij} = \beta_0 + \beta_1\sigma_i + \beta_2\alpha_i + \beta_3\lambda_i + \beta_4\rho_{ij} + \beta_5\rho_{ij}^2 + \beta_6 HR_i \\ + \beta_7(\sigma_i \times HR_i) + \beta_8(\alpha_i \times HR_i) + \beta_9(\lambda_i \times HR_i) + \phi\mathbf{X_i} + \theta_k + \varepsilon_{ij}. \end{aligned} \tag{1.9}$$

It is worth noting that we do cluster standard errors at the session level for this equation while we do not for other equations in the paper. It is because the effect of lower basis risk (HR) on index insurance demand *per se* is not the primary interest of this paper but a focus of our another paper and there is no reason to assume that risk parameters, the main focus of the study, would be correlated within session. Therefore, we cluster standard errors only when the difference in basis risk probabilities is considered to be important in predicting insurance demand in relation to risk attitude.

The results are shown in column 1 of Table 1.4. While none of the risk parameters conditional on HR is statistically significant, the signs of the coefficients on the risk parameters are consistent with our hypotheses. These results suggest that relative to the higher basis risk case (LR product), farmers who are risk averse or overweight a small probability of negative basis risk demand less insurance while loss averse farmers demand more insurance.

Heterogeneity Analysis with Insurance Price

In Section 1.2, we show that under reference point W, the marginal disutility of loss aversion becomes larger when an individual faces a higher premium only when his σ is less than 1. The rich demand data for each individual allows us to test for the heterogeneous effect of loss aversion by insurance prices. Because the maximum value of σ among the sample farmers is 0.9, we use a full sample.

Examining how the effect of loss aversion on WII demand varies by insurance premiums, we interact price variables and λ and add these two interactions to equation (1.7), which results in the following model.

$$\begin{aligned} Q_{ij} = \beta_0 + \beta_1\sigma_i + \beta_2\alpha_i + \beta_3\lambda_i + \beta_4\rho_{ij} + \beta_5\rho_{ij}^2 + \beta_6(\lambda_i \times \rho_{ij}) \\ + \beta_7(\lambda_i \times \rho_{ij}^2) + \beta_8 HR_i + \phi\mathbf{X_i} + \theta_k + \varepsilon_{ij}. \end{aligned} \tag{1.10}$$

Table 1.4, column 2 shows the results. The coefficient on the interaction term between insurance premium and loss aversion is negative and significant at the 10% level (p-value: 0.058). Consistent with our hypothesis, this result suggests that the negative effect of loss aversion on index insurance demand becomes greater with increasing insurance premium. We also find that the sign of the loss aversion variable interacted with the price squared is positive. Although it is marginal and insignificant, this term implies that the negative effect of loss aversion on insurance demand may not be linearly increasing with price but increasing at a decreasing rate.

The diminishing sensitivity to losses is another important element in PT, capturing individuals' decreasing marginal disutility to additional losses. Given that farmers were asked their demand for the same insurance product at different insurance prices sequentially, the only component affecting the size of losses is an insurance premium. Therefore, it is expected that loss averse farmers decrease insurance demand in response to a marginal increase in insurance premium but the negative marginal change in demand is decreasing with rising premiums. We can explore if this is the case by evaluating $\beta_6 + \beta_7 \times \rho$ at continuous insurance prices starting from the lowest unit price (50 KSH/unit).

In Figure 1.2, we report the marginal effect of loss aversion conditional on premium evaluated at different premium levels with its corresponding 95% confidence interval. In this graph, the x axis represents insurance premium and the y axis represents $\beta_6 + \beta_7 \times \rho$. We can clearly see that farmers have decreasing sensitivity to additional losses from increasing premiums.

Overall, our finding of the heterogeneous effects of loss aversion supports that the sample farmers' reference points are W as opposed to $W - \rho$ and therefore a pay-upfront insurance premium matters (Serfilippi et al., 2016; Casaburi and Willis, 2018) in the sample farmers' insurance purchase decisions, especially for those who have high degrees of loss aversion. Moreover, this finding suggests that farmers consider the pay-upfront premium as a "loss" not a "cost to reduce future uncertainty".

One can better understand farmers' sensitivity to premiums by considering the four outcomes under reference point W in Section 1.2. Farmers have one source of a gain due to positive basis risk and three sources of a loss. Based on the previous discussion of the positive basis risk event perceived by the farmers, the three loss outcomes are considerably more salient to the sample farmers than the single gain occasion. We also note that farmers are only able to control the premium among the parameters. As a result, those who are highly loss averse are less likely to purchase insurance when facing high premiums to minimize the salient domain where a loss is experienced.

1.6 Discussion and Concluding remarks

In this paper, we extend the scant knowledge of relations between risk parameters and index insurance demand. First, we find that prospect theory offers a better framework to explain farmers' WII choices than expected utility theory. We also find a positive effect of risk aversion along with a negative effect of loss aversion, suggesting that the inverse relationship between the risk aversion parameter (elicited based on an EU framework) and index insurance demand could be due to a confounding effect of risk and loss aversion parameters.

We also offer theoretical models based on prospect theory by assuming two distinctive reference points relative to which farmers evaluate gains and losses. The only difference between the reference points is an insurance premium. Based on the models, we examine which reference point is empirically supported. We find that farmers who are highly loss averse are less likely to buy index insurance and this negative effect becomes more prominent when farmers are offered the insurance products at high prices. Combined with theoretical predictions of index insurance participation, these results show our sample farmers do not include a pay-upfront insurance premium in their reference points, suggesting that they view the premium as a loss rather than a sunk-cost to reduce risk at harvest.

Why are these farmers sensitive to the pain of a loss from the premium, when purchasing insurance may prevent the larger pain of crop losses from the rainfall shocks? We first consider the fact that the sample farmers have limited experience with insurance. For them the need to pay a premium before a shock is realized, may not be straightforward[18] and they may have a narrow frame of WII. Also, their high liquidity constraints (Casaburi and Willis, 2018) may play a role in a way that charging premiums upfront reduces farmers' ability of consumption smoothing between the time of the premium payment (usually at planting) and the next harvest season. The data does not allow us to test which explanation addresses the question better, but our results suggest that more research is needed to better understand why paying a premium upfront is a loss for smallholder farmers. For instance, it would be interesting to see how a delay in premium payment affects WII demand for farmers with different degrees of loss aversion.[19]

Importantly, even though we were encouraged by farmers' high take-up elicited in the auction, their choices do not translate to actual purchases. Despite logistic reasons to dampen the actual take-up, this study shows that farmers would essentially have no actual willingness

[18]Dercon et al. (2014) provide a training session about index insurance to Ethiopian farmers in which they explain why their subjects need to pay premiums before knowing if rainfall shocks are realized.

[19]Casaburi and Willis (2018) shows that deducting an insurance premium from harvest revenues *at harvest* increased uptake of a crop insurance by 67 percentage points.

to purchase insurance. Therefore, our findings are limited to insurance demand choices in a hypothetical setting for which more research about risk parameters and index insurance demand based on actual data is needed.

By showing how smallholder farmers view a pay-upfront insurance premium, we can draw critical policy implications. Careful designs of insurance payment schemes to reduce a feeling of loss from the premium potentially make WII products appealing. For instance, a group-based index insurance plan, the basic concept of which is borrowed by micro-finance group lending, would be effective to increase WII demand (Dercon et al., 2014). Through sharing costs and benefits of WII, farmers decrease the pain of a loss due to an upfront payment of the insurance premium, which leads to an increase in index insurance uptake. Charging the premium through deductions from harvest revenues may effectively increase demand although enforcing delayed premium payments would be a big challenge (Casaburi and Willis, 2018).

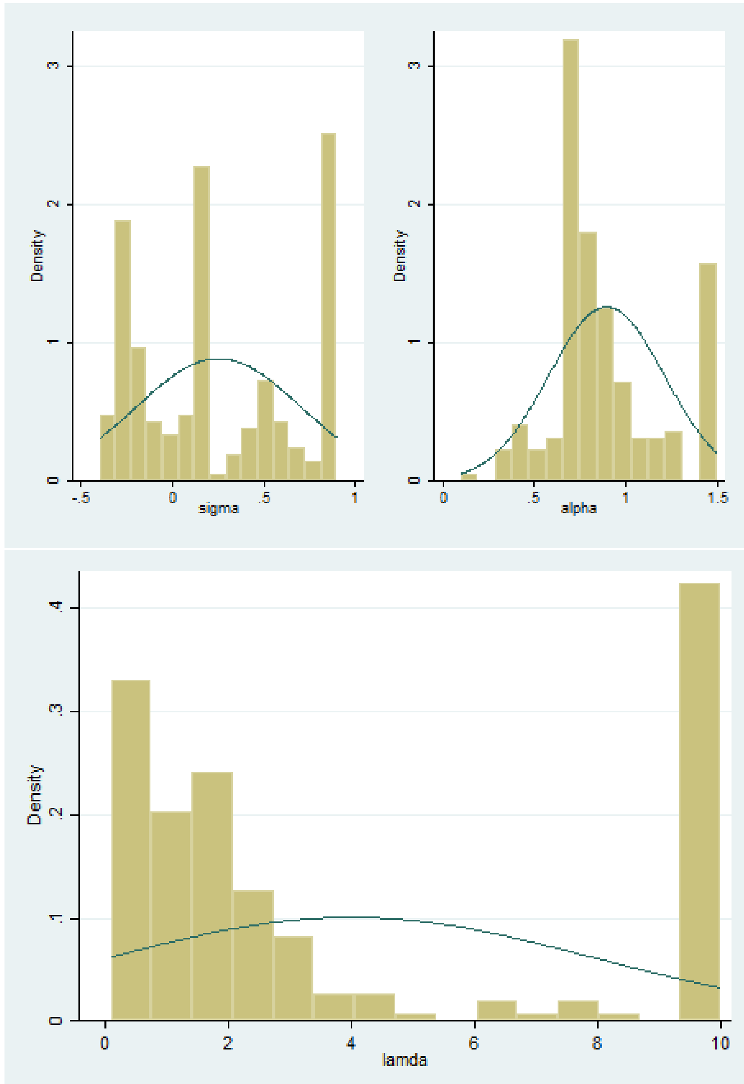


Figure 1.1: The Distributions of Risk Preference Parameters

Notes: This figure illustrates the distribution of risk preference parameters. The data shows that 64.4 percent of farmers are risk averse ($\sigma > 0$) and 66.5 percent exhibit loss aversion ($\lambda > 1$). Also, 69.9 percent of the sample overweights small probabilities and underweights large probabilities ($\alpha < 1$).

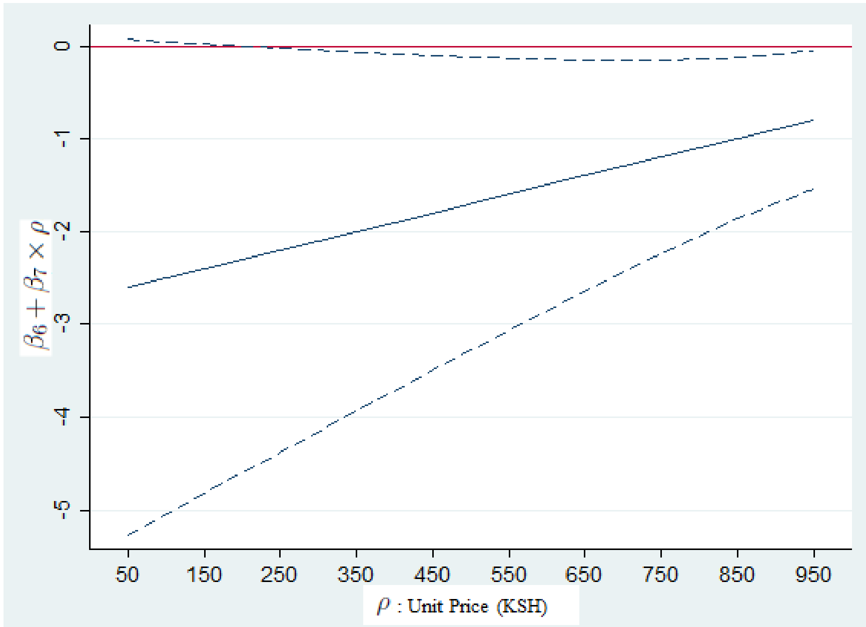


Figure 1.2: Average Marginal Effect of Loss Aversion conditional on Premium evaluated at Different Premium Levels with 95% CI

Notes: This figure shows that farmers have decreasing sensitivity to additional losses from increasing premiums.

Table 1.1: Summary Statistics of Sample Farmers

	Mean
Age	40.59
	(13.19)
Education	7.20
	(3.77)
Literacy	0.74
	(0.44)
Household head	0.66
	(0.47)
Male	0.26
	(0.44)
Land size (acre)	4.48
	(4.54)
Drought	2.81
	(1.10)
Loan availability	0.68
	(0.47)
Formal saving account	0.68
	(0.47)
Crop acre (acre)	1.91
	(1.10)
Net value (KSH)	51,503.34
	(69,558.86)
Sorghum	0.95
	(0.22)
Observations	239

Notes: This table presents the summary statistics for the sample we use. Literacy, Loan availability, and Formal saving account are indicators with a value of one if farmers said "yes". Crop acre represents acres of the insured crop and Net value is anticipated harvest values of the insured crop (1USD $\approx$ 100 KSH). Sorghum is an indicator for the insured crop being sorghum (i.e., crop types). Standard deviations are in parentheses.

Table 1.2: Prospect Theory Game Payoff Matrix (in KSH)

Lottery A		Lottery B	
Series1			
Card 1-7	**Card 8-10**	**Card 1-9**	**Card 10**
30	119	15	248
30	119	15	278
30	119	15	317
30	119	15	373
30	119	15	447
30	119	15	545
30	119	15	655
30	119	15	894
30	119	15	1191
30	119	15	1787
Series 2			
Card 1	**Card 2-10**	**Card 1-3**	**Card 4-10**
89	119	15	159
89	119	15	165
89	119	15	169
89	119	15	178
89	119	15	189
89	119	15	198
89	119	15	219
89	119	15	238
89	119	15	269
89	119	15	298
Series 3			
Card 1-5	**Card 6-10**	**Card 1-5**	**Card 6-10**
50	-8	59	-41
8	-8	59	-41
1	-8	59	-41
1	-8	59	-32
1	-16	59	-32
1	-16	59	-27
1	-16	59	-22

Table 1.3: Index Insurance Demand

	Q	Q
σ	2728.0**	2054.8
(risk aversion)	(1289.1)	(1261.5)
α	4677.2***	
(probability weighting)	(1595.6)	
D($\alpha < 1$)		-2209.8**
		(1074.7)
λ	-282.5*	-254.1*
(loss aversion)	(144.1)	(144.8)
ρ	-75.9***	-75.9***
(premium)	(5.64)	(5.64)
ρ^2	0.047***	0.047***
	(0.0052)	(0.0052)
HR	4980.8***	5039.3***
	(941.8)	(945.4)
Observations	1673	1673
R-squared	0.508	0.507

Notes: This table shows the effects of risk preference parameters and the difference in basis risk probabilities on index insurance demand, independently and interactively. The sample is top coded at the 99th percentile of demand from the whole sample of the project. Control variables include age, education, literacy, male, land size, drought, loan availability, formal saving, acres of crop, projected net value of production, crop types, and enumerator fixed effects. Standard errors are in parentheses with only the standard errors in column 3 being clustered at the session level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.4: Index Insurance Demand: Heterogeneous Effect of Risk Parameters

	Q	Q
σ	3011.2	2728.0**
(risk aversion)	(1902.4)	(1287.9)
α	2530.3	4677.2***
(probability weighting)	(4452.4)	(1594.0)
λ	-300.4	314.4
(loss aversion)	(284.4)	(305.8)
ρ	-75.9***	-65.1***
(premium)	(10.2)	(8.01)
ρ^2	0.047***	0.039***
	(0.0071)	(0.0074)
$\lambda \times \rho$		-2.70*
		(1.42)
$\lambda \times \rho^2$		0.0020
		(0.0013)
HR	1350.3	4980.8***
	(5816.7)	(940.9)
$\sigma \times$ HR	-648.6	
	(3520.8)	
$\alpha \times$ HR	3966.8	
	(6281.5)	
$\lambda \times$ HR	50.7	
	(422.2)	
Observations	1673	1673
R-squared	0.484	0.588

Notes: This table shows heterogeneous effects of risk parameters, focusing on non-linear probability weighting and loss aversion on index insurance demand. The first column shows how lowering a probability of negative basis risk (i.e., update to HR product from LR) affects index insurance demand through non-linear probability weighting and loss aversion parameters. The second column shows how the effect of loss aversion on index insurance demand varies by insurance premiums. The sample is top coded at the 99th percentile of demand from the whole sample of the project. Control variables include age, education, literacy, male, land size, drought, loan availability, formal saving, acres of crop, projected net value of production, crop types, and enumerator fixed effects. Standard errors are in parentheses with only the standard errors in column 1 being clustered at the session level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Chapter 2

Panel price index with price, quality, and variety effects: using chinese scanner data on non-alcoholic beverages

2.1 introduction

Cost-of-living indexes are critical to measuring real income and expenditures of household living in different locations. These indexes are fundamental to welfare analysis because household real income and expenditures are used to estimate degrees of poverty and income inequality. Despite their importance, the standard cost-of-living indexes may not capture actual price variations because they often partially address product heterogeneity such as quality and variety. For instance, the U.S. Consumer Price Index (CPI) accounts for product entry and exit by only examining particular surveyed stores although product availability varies across regions (Handbury and Weinstein, 2015) and over time (Broda and Weinstein, 2010). Previous studies have shown that either quality or variety plays an important role in constructing price indexes or measuring welfare effects, for instance, from trade (Bils and Klenow, 2001; Broda and Weinstein, 2010; Khandelwal, 2010; Sheu, 2014; Feenstra et al., 2017).

This challenge is particularly severe in large developing countries such as China where comprehensive expenditure data is not readily available[1] and spatial heterogeneities in consumer preferences and product availability are likely to be substantial. China does not have

[1] Deaton and Dupriez (2011) have no choice but to use unit values of foods to construct spatial price indexes for rural and urban areas of India and Brazil, due to inherent limitations of household consumption survey data. Unit value embodies both the quality of purchases and market price variation. They show quality unadjusted unit values could bias the price indexes and overstate the urban/rural price difference.

official price indexes that reveal cross-regional price differences. Responding to this knowledge gap, Brandt and Holz (2006) generate spatial deflators to measure inequality in China by exploiting regional price data published in the Chinese Statistical Yearbooks. Using data from the Chinese household income project with other supplementary data, Almås and Johnsen (2012) construct spatial price indexes using Engel's law and apply them to measure poverty. Due to data limitations, however, these studies are unable to account for product heterogeneity. One study (Feenstra et al., 2017) on cost-of-living indexes in China addresses this limitation by using detailed barcode level data but is limited to considering variety only. To our knowledge, no previous study has accounted for product quality in measuring cost-of-living in China.[2] Yet, the studies mentioned earlier provide evidence that a failure to address variety itself could largely bias the true cost-of-living indexes across locations, especially in large countries.

The purpose of this study is to show the need to account for the two dimensions of product heterogeneity in cost-of-living indexes by quantifying product heterogeneity biases. To this end, we first build panel price indexes that account for the heterogeneity for 69 markets (a combination of a province and a year) using a large-scale retail scanner data of Chinese non-alcoholic beverages (NB) collected from urban consumers. We then decompose these price indexes into three components (i.e., price, quality, and variety) to determine which component contributes the most to cost-of-living differences. To measure biases arising from ignoring product quality and variety, we compare our price indexes to superlative price indexes such as Fisher ideal and Törnqvist that use matched-basket methods in which quality and variety are not addressed.

We find that neglecting product variety and quality effects in cost-of-living indexes may result in underestimates of welfare gaps between two markets. The magnitude of the downward bias in a measured welfare gap is up to 94.5% when we compare our price indexes to the two superlative indexes. Additionally, the product heterogeneity bias is relatively more

[2]There is a handful of studies that investigates the implications of quality change in the context of the United States (Bils and Klenow, 2001; Bils, 2004; Broda and Weinstein, 2010).

evident in the spatial dimension than in the temporal dimension. Our price indexes indicate that the spatial welfare gap between the base and a comparison market is particularly considerable when the comparison market is a major large city such as Beijing and Shanghai, while the two superlative indexes indicate there is little difference in across-province cost-of-living. Because the base market of the price indexes, which is equivalent across all markets, has the smallest product varieties with the lowest economic development (in terms of per capita GDP), the stark welfare differentials between the base market and large cities shown in our panel price indexes are not surprising. This finding is also consistent with the evidence in Feenstra et al. (2017) that the cost of living for grocery-store products is lower in larger cities.[3] Based on these facts, our results imply that the product heterogeneity is important in measuring cost-of-living differences, especially in the spatial dimension.

The price decomposition results show that product quality and variety are the primary sources of the differences in cost-of-living for NB between two markets. Indeed, the values of the price component in the decompositions are similar across markets. We also find that the magnitude of the price component is mostly analogous to that of the two superlative price indexes for all markets. Given that the price component represents the distribution of between-market price differences while the superlative price indexes only compare prices of *identical* products between markets, this finding suggests that the price distributions are similar across markets irrespective of product combinations in the fixed baskets, weighting more on the important roles of the product quality and variety effects in determining the cost-of-living indexes for NB.

We also find that the directions of quality and variety effects on the price indexes are the same for provinces with major large cities in a way that consumers in those provinces enjoy broader variety and higher quality. However, the directions are opposite to each other for provinces with small and medium-size cities. The latter case implies that a broader choice

[3]However, while Feenstra et al. (2017) found the sources of the lower price level in larger cities from their greater variety and lower prices (quality is not accounted for in this study), our study found the sources from their higher quality and broader variety.

set of products is not necessarily associated with higher quality. This situation is possible in the Chinese non-alcoholic beverage market because many local brands produce a wide range of products that are only sold on local markets. If the quality of these products do not exceed that of national/international brands, we expect to see the effects of variety and quality cancel out in cost-of-living indexes.

Besides our novel findings, this study provides an important methodological contribution. Given that previous studies to date have focused on either temporal or spatial differences (see, for example, Jolliffe, 2006; Albouy, 2009; Broda and Weinstein, 2010), we contribute to a knowledge gap of the literature by building panel price indexes. In constructing panel price indexes, we set one market of which the product variety is the smallest as the base market. This approach has the advantage in that it allows one to compare price levels of non-alcoholic beverages across regions in a given period as well as those within a province over time. Moreover, this setting enables us to apply the unique sampling method of Sheu (2014) in which the price index is decomposed into price, quality, and variety components that are consistent across all markets . Using this method, we disentangle the portions of price indexes by which price and quality contribute to, respectively, from the portion of the variety term. That is, for each comparison market we sample the same number of products as the base market but from the distribution of price and quality of the corresponding comparison market to capture any differences in the price and quality of the comparison market from those of the base while holding variety constant at the base market.

We also fully exploit our panel data. First, we explicitly quantify unobserved (by econometricians) time-invariant product quality through barcode fixed effects. This approach is an extension of the insight of measuring quality change that is developed in the literature on price indexes and trade. The implication is that conditional on price, products with a higher quality will result in a higher market share through a lower quality-adjusted price (Khandelwal, 2010; Broda and Weinstein, 2010). Under this insight, previous studies, mostly focusing on temporal price variations (i.e., inflation), implicitly address quality change by specifying the price indexes only in terms of prices and market shares. Our approach, instead, estimates the value of each product quality from a demand equation in which the same insight is embedded and measures the effect of quality differences on the cost-of-living variations that is disentangled from the effects of price and variety. Second, using within transformation and instrument variables derived from the panel structure, our study largely controls for endogeneity bias of the demand equation. Finally, we can clearly identify demand parameters including prices and elasticities of substitution using sufficient variations in detailed product-level data at relatively high frequency across regions and over time.

The organization of this paper is as follows. Section 2.2 provides data overview. In Section 2.3, we describe theoretical models for consumer utility, price indexes, and price index decompositions. Section 2.4 presents demand models estimations and a method to correct a sample size bias. Section 2.5 outlines the empirical results. Section 2.6 concludes the paper.

2.2 DATA OVERVIEW

The barcode-level non-alcoholic beverages purchasing dataset is taken from Kantar Worldpanel database. These data were collected by Kantar from a sample of 39,000 households in 23 Chinese provinces[4] from 2011 to 2013. Similar to Nielsen HomeScan data, households participating in the Kantar's household scanner data program are asked to scan in every beverage purchase they made using handheld Universal Product Code (UPC, or barcode) scanners. Households earn points in return for their participation that can be used to redeem products that the Worldpanel does not track. This reward system was designed to not generate confounding factors interfering with households' purchase decisions on products in interests. Although one could argue that home-scan data may underestimate consumers' actual purchases, this approach has shown to be effective for measuring purchasing patterns

[4] The 23 provinces consist of three municipalities and 20 provinces which are part of the first tier of administrative divisions in China directly under the central government.

at the household level (Handbury and Weinstein, 2015), especially given that no alternative exists to obtain such detailed micro data.

Kantar also collects household demographic information to calculate sampling weights. The sampling weights are applied to the UPC purchasing data to generate demographically representative purchases. Then, the purchasing records are aggregated into four income groups.[5] As a result, each observation in our data, the purchase of an individual UPC, is represented by following two variables; weighted expenditure in the Chinese national currency (RMB) and weighted volume by liters, both of which were reported for every 4-week period and aggregated at the income group levels. Focusing on product quality and variety, we integrated four income categories into one.

We obtain a quad-week level price of each barcode product j as RMB/liter by dividing its total weighted expenditure by its total weighted volume within each region and quad-week period. Accordingly, the price of each UPC is same for every household living in the same provinces at the same time periods, which is consistent with a theoretical utility model described in subsection 2.3.1. When we simply divide the weighted expenditure by the weighted volume for each observation in our data, households from different income groups face different prices for the same good j in region c at period p. These discrepancies can exist when households with unobserved specific characteristics are more likely to use price-cutting strategies available to them such as coupons and vouchers, compared to households without these characteristics. When considering the method by which we calculate prices, these unobserved household characteristics imply that we need to address endogeneity in prices. More discussion related to this issue is in subsection 2.4.1.

The non-alcoholic beverage products are categorized into seven sub-groups: Carbonated soft drinks (CSD), Juice, Ready-to-drink (RTD) tea, Functional drinks, Packaged water, RTD coffee, and Soybean milk. We use three subcategories (i.e., CSD, Juice and RTD tea) among seven that allow us to have one common base market in terms of spatial and time

[5]The four different income groups are defined as follows: monthly income less than 3,000 RMB, 3,001-5,000 RMB, 5,001-7,000 RMB, and over 7,000 RMB.

dimensions.[6] Table 2.1 shows summary statistics for the sample we use to estimate demand. The number of unique UPC varies widely across provinces and periods providing sufficient variations to identify demand parameters.

2.3 MODEL

In this section, we discuss theoretical frameworks that form the panel price indexes and price decompositions. We specify a consumer demand model using modified logit preferences in order for the nested logit (NL) model to be aligned with the nested constant elasticity of substitution (NCES) model. The specific models are the modified nested logit (hereafter, NL-NCES).[7] In the model, products within a sub-group are more similar substitutes compared to products between sub-groups. For instance, juice product A is a closer substitute for juice product B compared to an RTD tea product.

2.3.1 Consumption Utility

We first explain the consumption utility underlying the NL-NCES model. The utility for household i who purchases product j in region c at period p is

$$u_{ijcp} = \ln(a_{jc} m_{ijcp}) + \xi_{igcp} + \epsilon_{ijcp} \tag{2.1}$$

where a_{jc} is observed and unobserved (by econometricians) product-specific quality and, m_{ijcp} is the quantity of product j in region c that household i chooses to buy. ξ_{igcp} and ϵ_{ijcp} describe *i.i.d.* random draws from a logit distribution with scales μ_1 and μ_2, respectively: one for each group $g \in \{\text{CSD, JUICE, TEA}\}$ and the other for each product $j \in B_{gcp}$ (the bundle of goods in group g purchased by consumers in region c at period p). The last two terms fall into the error terms.

[6] The three groups account for 80.4% of total expenditure and 70.28% of total volume of the data we observe.

[7] By choosing the modified logit framework, we avoid a computational burden in demand estimations such as a non-convergence of the numerical search that often occurs in the case of the normal distribution (see, Knittel and Metaxoglou, 2014).

It is worth noting that our utility model differs slightly from the standard logit model (Nevo, 2000a) in that every utility component except for the random part enters in logarithms as opposed to levels. This modification is necessary for the nested logit (NL) model to be in line with the NCES, as an extension of Anderson et al. (1992)'s work, which provides the following two benefits. First, we can assume a relatively more realistic purchasing pattern where consumers are allowed to purchase more than one unit of a chosen beverage good that gives them the maximum utility. Second, the strong theoretical tie between two models allows us to build our price indexes and price index decompositions in an NCES structure using parameters that are estimated from a demand model based on the NL-NCES utility function.

Substituting the budget constraint, $y = P_{jcp} m_{ijcp}$, we can obtain indirect utility for the household as:

$$\nu_{ijcp} = \ln a_{jc} - \ln P_{jcp} + \ln y + \xi_{igcp} + \epsilon_{ijcp} \tag{2.2}$$

We assume that a_{jc} varies across regions but it is fixed over time. This is because obvious changes in product characteristics should cause new barcodes to be assigned (GTIN Allocation Rules, 2007) and three years are relatively short for dramatic market changes to occur. The quality term, a_{jc}, therefore, consists of the national mean valuation of the quality of product j (e.g., 100% orange juice vs. orange-flavored juice, national brand premiums over generic brand, any brand equity) and regional market specific unobserved characteristics such as impacts of promotional activities or unobserved demand shocks.

We obtain the following expenditure shares by 1) integrating over the remaining logit random shocks (i.e., the last two error components in equation (2.2)) and 2) converting the resulting conditional probabilities of choosing good j and group g to expenditure shares.

Note that we can derive the same expenditure share from an NCES utility structure of a representative consumer due to the relationships between NL and NCES.[8]

$$s_{jcp} = \frac{b_{jc} {P_{jcp}}^{1-\sigma}}{\left(\sum_{k \in B_{gcp}} b_{kc} {P_{kcp}}^{1-\sigma}\right)^{\frac{\gamma-\sigma}{1-\sigma}} \sum_{g \in \{\text{CSD, JUICE, TEA}\}} \left(\sum_{k \in B_{gcp}} b_{kc} {P_{kcp}}^{1-\sigma}\right)^{\frac{1-\gamma}{1-\sigma}}} \tag{2.3}$$

where $b_{jc}(= {a_{jc}}^{\frac{1}{\mu_2}})$ is a measure of quality, $\sigma = (1 + \frac{1}{\mu_2})$ is the elasticity of substitution between beverages in the same group, and $\gamma = (1+\frac{1}{\mu_1})$ is the elasticity of substitution between groups. We derive an estimating demand equation from (2.3). Importantly, the expenditure share increases as a measure of quality becomes larger given all else is being constant. As mentioned in the introduction, previous studies use this relationship to implicitly address quality in price indexes in that they express the price indexes only in terms of the expenditure share and prices. We also rely on this relationship to understand how product quality mediates pathways between demand and our price indexes. Unlike the previous studies, however, we explicitly estimate the quality parameter, b_{jc}, to examine how important quality differences are in determining the cost-of-living for non-alcoholic beverages. We discuss the estimation of the quality measurement in detail in Section 2.4.

2.3.2 Price Indexes

We build the price indexes at the 'region-year' levels. Therefore, we define a market as a combination of a region and a year and denote it by m. Our key strategy in constructing comparable panel price indexes is to set an equivalent base market where the number of unique barcode products is the smallest for all markets. In this context, one can simply draw

[8]The corresponding utility function is given by :

$$u_{cp} = \Big(\sum_{g \in \{\text{CSD, JUICE, TEA}\}} \Big(\sum_{j \in B_{gcp}} {b_{jc}}^{\frac{1}{\sigma}} {m_{jcp}}^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma-1}{\sigma}\frac{\gamma-1}{\gamma}} \Big)^{\frac{\gamma}{\gamma-1}}$$

a bilateral comparison of the cost-of-living indexes using any markets by dividing one by the other. In other words, transitivity is satisfied (Hill, 1997; Diewert, 1999).

Suppose we choose one base market that has the least number of product varieties in terms of 1) the broad category, non-alcoholic beverages, and 2) the three sub-groups of the beverages: $g \in \{\text{CSD, JUICE, TEA}\}$. Denote the base market by m^0, the combination of region c^0 and year y^0, and the comparison market by m^1 in the same manner.

In each market, consumers have various bundles of non-alcoholic beverage products (B_m) available to purchase. Each unique barcoded product j in market m is differentiated by its price P_{jm} and quality b_{jm} from which consumers obtain utility. The elasticity of substitution within the sub-groups (σ) is assumed to be greater than the elasticity of substitution between the sub-groups (γ). In this setting, the price index for market m^1 (π_{m^1}) is given by

$$\pi_{m^1} = \frac{\left(\sum_{g \in \{\text{CSD, JUICE, TEA}\}} \left(\sum_{j \in B_{gm^1}} b_{jm^1} {P_{jm^1}}^{1-\sigma} \right)^{\frac{1-\gamma}{1-\sigma}} \right)^{\frac{1}{1-\gamma}}}{\left(\sum_{g \in \{\text{CSD, JUICE, TEA}\}} \left(\sum_{j \in B_{gm^0}} b_{jm^0} {P_{jm^0}}^{1-\sigma} \right)^{\frac{1-\gamma}{1-\sigma}} \right)^{\frac{1}{1-\gamma}}} \tag{2.4}$$

Note that π_{m^1} is a ratio of two NCES unit expenditure functions, one for the base and the other for the comparison market. Based on the perfect convertibility between NL and NCES, the price indexes are termed NL-NCES price indexes.

We interpret the price index according to the magnitude by which prices of a chosen base market with lower (higher) levels of variety/quality would have to decrease (increase) in order to achieve the same level of welfare of consumers from a comparison market with higher (lower) variety/quality.

We find that "Gui Zhou-2011" market has the smallest number of unique products in terms of the overall category *and* all sub-groups. Defining this market as the base market enables us to build the price indexes as well as the price index decompositions which need a sampling method requiring one common base market that has the smallest varieties for all sub-groups. We further discuss the sampling method and the price index decomposition formula in the next subsection.

2.3.3 Price Index Decompositions

Derivations

We show how our NL-NCES price index is decomposed into price, quality, and variety terms in an NCES structure. The main point of the decomposition is to disentangle the variety effect from the price and quality effects. This can be done by 1) identifying identical products between two markets and 2) generating a variety-adjustment term based on the seminal work by Feenstra (1994).

The NL-NCES price index consists of two tiers. The upper tier (indexed as superscript U) includes three sub-groups (CSD, JUICE, and TEA) and the lower tier (indexed as superscript L) contains the bundle of products within each sub-group. To obtain overall price index decompositions, we first construct the price index decompositions for each tier and then combine these two tier-specific decompositions. We demonstrate a derivation of the price decompositions for the upper tier only because the same derivation is applied to the lower tier.

We define a CES utility function of a representative consumer for the upper tier over the consumption (M_{gm}) of each "product[9]" g in market m^1 with its quality, ${b^U}_{gm^1}$:

$$ {u^U}_{m^1} = [\sum_{g \in B^U_{gm^1}} {b^U}_{gm^1}{}^{\frac{1}{\gamma}} M_{gm^1}{}^{\frac{\gamma-1}{\gamma}}]^{\frac{\gamma}{\gamma-1}} \tag{2.5} $$

where γ is the elasticity of substitution between products g in the upper tier.

Based on (2.5), the upper tier unit expenditure function defined over the price (P_{gm^1}) of each product g in market m^1 is given by:

$$ {C^U}_{m^1} = [\sum_{g \in B^U_{gm^1}} {b^U}_{gm^1} P_{gm^1}{}^{\frac{1}{\gamma}}]^{\frac{1}{1-\gamma}} \tag{2.6} $$

We obtain the expenditure share for each product g in market m^1 by applying Shephard's Lemma to formula (2.6).

$$ {S^U}_{gm^1} = \frac{P^{1-\gamma}_{gm^1} b^U_{gm^1}}{\sum_{g \in B^U_{m^1}} P^{1-\gamma}_{gm^1} b^U_{gm^1}} = \frac{P_{gm^1} M_{gm^1}}{\sum_{g \in B^U_{m^1}} P_{gm^1} M_{gm^1}} \tag{2.7} $$

[9]In the upper tier, each non-alcoholic beverage group is considered as a product.

Analogously, we obtain the expenditure share in the base market m^0. Following the definition of the variety-adjustment term shown in Feenstra (1994), we identify a bundle of common products ($B^U_{m^0m^1} = \{\text{CSD, JUICE, TEA}\}$) available in both markets and generate the variety-adjustment terms $\lambda^U_{m^0m^1}$ and $\lambda^U_{m^1m^0}$ that are defined as below:

$$\lambda^U_{m^0m^1} = \frac{\sum_{g\in B^U_{m^0m^1}} P_{gm^0} M_{gm^0}}{\sum_{g\in B^U_{m^0}} P_{gm^0} M_{gm^0}}, \quad \lambda^U_{m^1m^0} = \frac{\sum_{g\in B^U_{m^0m^1}} P_{gm^1} M_{gm^1}}{\sum_{g\in B^U_{m^1}} P_{gm^1} M_{gm^1}} \tag{2.8}$$

Multiplying the numerator and denominator of the farthest right-hand side in equation (2.7) by $\sum_{g\in B^U_{m^0m^1}} P_{gm^0} M_{gm^0}$ and $\sum_{g\in B^U_{m^0m^1}} P_{gm^1} M_{gm^1}$, and applying the second variety-adjustment term, $\lambda^U_{m^1m^0}$ in (2.8), we can express the expenditure share, (2.7), in terms of expenditure share for common goods and a variety term. We apply this computation to the expenditure share in the base market as well, which results in:

$$S^U_{gm^0} = \lambda^U_{m^0m^1} S^{*U}_{gm^0}, \qquad S^U_{gm^1} = \lambda^U_{m^1m^0} S^{*U}_{gm^1} \tag{2.9}$$

where $S^{*U}_{gm^0} = \frac{P_{gm^0} M_{gm^0}}{\sum_{g\in B^U_{m^0m^1}} P_{gm^0} M_{gm^0}}$ and $S^{*U}_{gm^1} = \frac{P_{gm^1} M_{gm^1}}{\sum_{g\in B^U_{m^0m^1}} P_{gm^1} M_{gm^1}}$ are the shares of each common product g in expenditure for the upper tier that are defined by its price and consumption in market m^0 and m^1, respectively.

We then decompose the upper tier unit expenditure functions for both markets into the price, quality, and variety terms along with $S^{*U}_{gm^0}$ and $S^{*U}_{gm^1}$, respectively by exploiting the relationship between (2.6) and (2.7) as well as the relationship between $S^U_{gm^1} = \frac{P^{1-\gamma}_{gm^1} b^U_{gm^1}}{{C^U_{m^1}}^{1-\gamma}}$,the resulting term from (2.6) and (2.7), and (2.9).

$$C^U_{m^0} = \left(\frac{1}{\lambda^U_{m^0m^1}}\right)^{\frac{1}{1-\gamma}} P_{gm^0} {b^U_{gm^0}}^{\frac{1}{1-\gamma}} \left(\frac{1}{S^{*U}_{gm^0}}\right)^{\frac{1}{1-\gamma}}, \quad C^U_{m^1} = \left(\frac{1}{\lambda^U_{m^1m^0}}\right)^{\frac{1}{1-\gamma}} P_{gm^1} {b^U_{gm^1}}^{\frac{1}{1-\gamma}} \left(\frac{1}{S^{*U}_{gm^1}}\right)^{\frac{1}{1-\gamma}} \tag{2.10}$$

We express the change in the cost-of-living from the base market m^0 to the comparison market m^1 as the ratio between the decompositions (2.10) for the two markets.

$$\frac{C^U_{m^1}}{C^U_{m^0}} = \left(\frac{\lambda^U_{m^0m^1}}{\lambda^U_{m^1m^0}}\right)^{\frac{1}{1-\gamma}} \left(\frac{P_{gm^1}}{P_{gm^0}}\right) \left(\frac{b^U_{gm^1}}{b^U_{gm^0}}\right)^{\frac{1}{1-\gamma}} \left(\frac{S^{*U}_{gm^0}}{S^{*U}_{gm^1}}\right)^{\frac{1}{1-\gamma}} \tag{2.11}$$

Lastly, taking a geometric mean across all common goods using normalized versions of the logarithmic means of cost shares (Sato, 1976; Vartia, 1976) as weights (ω_{gm^1}) we obtain the following price index decompositions for the upper tier.

$$\pi^U_{m^1} = \frac{C^U_{m^1}}{C^U_{m^0}} = \left(\frac{\lambda^U_{m^0m^1}}{\lambda^U_{m^1m^0}}\right)^{\frac{\omega_{gm^1}}{1-\gamma}} \prod_{g\in B^U_{m^0m^1}} \left(\frac{P_{gm^1}}{P_{gm^0}}\right)^{\omega_{gm^1}} \prod_{g\in B^U_{m^0m^1}} \left(\frac{b^U_{gm^1}}{b^U_{gm^0}}\right)^{\frac{\omega_{gm^1}}{1-\gamma}} \prod_{g\in B^U_{m^0m^1}} \left(\frac{S^{*U}_{gm^0}}{S^{*U}_{gm^1}}\right)^{\frac{\omega_{gm^1}}{1-\gamma}}$$

$$\text{where} \quad \omega_{gm^1} = \frac{\left(\frac{S^{*U}_{gm^1} - S^{*U}_{gm^0}}{lnS^{*U}_{gm^1} - lnS^{*U}_{gm^0}}\right)}{\sum_{g\in B^U_{m^0m^1}}\left(\frac{S^{*U}_{gm^1} - S^{*U}_{gm^0}}{lnS^{*U}_{gm^1} - lnS^{*U}_{gm^0}}\right)}. \tag{2.12}$$

The ratio of the variety-adjust terms is one when both markets have three upper tier "products" (i.e., CSD, JUICE and TEA) with no "product" turnover. Also, the last product term in (2.12) equals unity, which can be easily shown when we take its natural log.

Similarly, we can obtain the price index decompositions ($\pi^L_{m^1}$) for the lower tier where consumer utility functions are defined over product j in each bundle of three sub-groups of beverages (i.e., $j \in B^L_g$ where $g = \{\text{CSD, JUICE, TEA}\}$). We then combine the price index decompositions of both tiers into one by substituting $\pi^L_{m^1}$ into the upper tier price ratio terms and assuming no difference in quality parameters of the upper tier "products" between two markets. As a result, the overall NL-NCES price decompositions are given by:

$$\pi_{m^1} = \prod_{g\in B^U_{m^0m^1}} \left[\left(\prod_{j\in B^L_{g,m^0m^1}} \left(\frac{P_{jm^1}}{P_{jm^0}}\right)^{\omega_{jgm^1}}\right)^{\omega_{gm^1}} \left(\prod_{j\in B^L_{g,m^0m^1}} \left(\frac{b^L_{jm^1}}{b^L_{jm^0}}\right)^{\frac{\omega_{jgm^1}}{1-\sigma}}\right)^{\omega_{gm^1}} \left(\frac{\lambda^L_{m^0m^1}}{\lambda^L_{m^1m^0}}\right)^{\frac{\omega_{gm^1}}{1-\sigma}}\right]$$

$$\text{where} \quad \omega_{jgm^1} = \frac{\left(\frac{S^{*L}_{jgm^1} - S^{*L}_{jgm^0}}{lnS^{*L}_{jgm^1} - lnS^{*L}_{jgm^0}}\right)}{\sum_{g\in B^L_{g,m^0m^1}}\left(\frac{S^{*L}_{jgm^1} - S^{*L}_{jgm^0}}{lnS^{*L}_{jgm^1} - lnS^{*L}_{jgm^0}}\right)}, \text{and} \quad B^U_{m^0m^1} \in \{\text{CSD, JUICE, TEA}\}. \tag{2.13}$$

where σ is the elasticity of substitution within sub-group products in the lower tier.

Common products sampling

Given NL-NCES price decompositions formula (2.13), we need to generate a bundle of common products observed in both markets for each sub-group, indexed by B^L_{g,m^0m^1}. Unlike Feenstra (1994) where B^L_{g,m^0m^1} contains the actual common products in both markets, we create synthetic samples for this bundle to capture overall changes in price and quality between two markets. We use the novel sampling methodology developed by Sheu (2014).

We first limit the sample size of each sub-group in a comparison market to the number of unique barcode products of the corresponding subgroup in the base market. We then draw a comparison market sample in a way that their prices and quality represent the distribution of these two components in the corresponding market. This is easiest to understand in the context of an example. Suppose that the base market has A different products available in CSD bundle and a chosen comparison market has B differentiated products where B≥A. We split the list of price and quality of CSD products in the comparison market into five cross percentile bands, generating 25 price/quality bins. From each bin, we randomly select products at a rate of A/B, thereby obtaining a CSD bundle for the comparison market that is the same size as the CSD bundle for the base market.

We order the products in each bundle of both markets by increasing quality[10] and match these sorted products in one market to those in the other to create the bundle of common CSD products (B^L_{CSD,m^0m^1}). We then calculate price index decompositions according to (2.13) for CSD. After this sampling process is completed for JUICE and TEA, values of price, quality, and variety terms in the price index decomposition are generated. We repeat this procedure

[10]The resulting decompositions have similar quantitative interpretations when we order the products by increasing prices, which is consistent with the theory. Therefore, the fact that the results are not affected by the way of sorting provides evidence that the sorted products that are matched one another do not change that much (i.e., Products with low price mostly have low quality) and that our results are robust. The former evidence is consistent with what economists normally believe: price is a signal of quality. However, by separating out quality and price effects on market share through demand estimation, we found that the between-market difference in the quality distributions is much wider than the difference in the price distributions (see, Section 2.5).

100 times to prevent biased outcomes that are possibly driven from one particular sample. We then obtain one value for each decomposition component by averaging over the 100 results.

This unique sampling method is particularly useful because we can separate the portions of price indexes by which price and quality contribute to, respectively, from the portion of variety terms while capturing overall variations in the first two components between the markets. That is, for each comparison market, we hold variety constant at the base market but capture any differences in the price and quality of the comparison market from those of the base.

Interpretation

Here, we provide an interpretation of each component of price decomposition (2.13). The geometric average of price ratios ($\frac{P_{jm^1}}{P_{jm^0}}$) captures the variations in prices between comparison and base markets, and the geometric average of quality ratios ($\frac{b^L_{jm^1}}{b^L_{jm^0}}$) contains information about quality differences. The last variety term captures the extent of product entry and exit in terms of the expenditure shares between two markets. Therefore, beyond the overall interpretation of price index in subsection 2.3.2, we can determine the respective portion of price, quality, and variety effects in the cost-of-living index that contributes to a change in welfare between two markets.

2.4 Estimation

2.4.1 Demand Model Estimation

The observation unit for estimation is indexed by region c and period p, following the consumption utility in subsection 2.3.1.[11] We do not aggregate cross-sectional observations by

[11]Therefore, a panel individual is defined as a product in a region. Under this definition, product A in region 1 is considered as a different product from product A in region 2, while it is to be the same product within region 1 over periods.

year, as what we did for the price index and decompositions, because aggregation biases are likely to occur due to significant information loss in the data.

We first choose a reference good within the non-alcoholic beverages (i.e., inside good) to rescale estimation inputs to be in units relative to those of the reference good. Our choice of the reference good differs from other studies based on standard logit demands where an outside good or no purchase option is defined as a reference good. This difference is from the perfect convertibility between the NCES and NL models. In the NCES model, expenditure of products is allocated under the condition that a given amount of income is labeled for the bundle of products in questions. In this context, any shifts of expenditure can be absorbed across products within nests of the model. Therefore, the reference good should be an inside good that appears in every region and period. We have only one product satisfying this criterion, which is a functional drink product (i.e., Nong-Fu-Shan-Quan fiber drink 550ml).

We obtain the demand model by taking the log of the expenditure shares formula (2.3) for each product j and the reference good and subtracting the latter from the former. We assume that the quality measure for the reference good, indexed by b_{0c}, equals one for all regions. We then express the demand equation as:

$$\ln\left(\frac{s_{jcp}}{s_{0cp}}\right) = \frac{\gamma-1}{\sigma-1}\ln b_{jc} - (\gamma-1)\ln\left(\frac{P_{jcp}}{P_{0cp}}\right) + \frac{\sigma-\gamma}{\sigma-1}\ln s_{jcp|g} \tag{2.14}$$

where s_{jcp} and s_{0cp} are the share of expenditure for product j and the reference good amongst consumers in region c at period p, respectively. $s_{jcp|g}$ is the share of expenditure for the product j within the product's corresponding group g, implying that $s_{0cp|g}$ equals one.

We apply panel fixed effects (FE) to estimate main parameters of the demand and recover estimates of the time-invariant fixed effects for b_{jc}. This is an extension of Nevo (2001)'s work advocating the use of product's brand fixed effects, focusing on variations between brands within a group in estimating ready-to-eat cereal demand at brand levels. This method allows price and quality effects on market shares to be separated out. A common way to capture the product-specific measure of quality is to use observed product characteristics. However, the choice of product characteristics tends to be arbitrary and may not effectively capture time-invariant differences across goods in different provinces. The equation for b_{jc} is assumed to be as:

$$\frac{\gamma - 1}{\sigma - 1} \ln b_{jc} = \frac{\gamma - 1}{\sigma - 1} k_{jc} \tag{2.15}$$

where the second term is the time-invariant fixed effects.

Combining equation (2.14) and (2.15), we can specify an estimating demand equation as:

$$\ln\left(\frac{s_{jcp}}{s_{0cp}}\right) = -(\gamma - 1)\ln\left(\frac{P_{jcp}}{P_{0cp}}\right) + \frac{\sigma - \gamma}{\sigma - 1}\ln s_{jcp|g} + \frac{\gamma - 1}{\sigma - 1}k_{jc} + \epsilon_{jcp} \tag{2.16}$$

We next address endogeneity problems in the demand estimation that are caused by unobservable components that fall into the error terms. Differentiated-products pricing models assume that prices are a function of marginal cost and a mark-up term. The mark-up term is a function of the unobserved product characteristics of the products, which induces endogeneity in prices. The unobserved product characteristics can also introduce endogeneity in the within group share variable ($\ln s_{jcp|g}$) because higher unobserved quality can lead higher within group sales.

We control for a significant amount of endogeneity in the log within group share variable with barcode fixed effects. The log price variable, however, still suffers from endogeneity due to how we obtain prices (Section 2.2). As noted by Nevo (2000a), using average prices introduces another endogeneity due to measurement error bias.

We, therefore, exploit the panel structure of the data (Hausman et al., 1994; Nevo, 2000b, 2001) to address the endogeneity in the price variable. Specifically, for product j in a region c of period p, we generate a regional average price of product j in the same period p excluding the region c. Our identifying assumptions are two fold. It is expected that the common marginal cost of product j causes prices of the product across provinces to be correlated. However, the region-specific valuations of the product are independent across provinces but are allowed to be correlated within a province over time once we control for product-specific fixed effects. Based on these two assumptions, we argue that the prices of the same barcode product in other provinces are valid instruments.

If there is no product j across other regions in the same period, we instrument this price by itself in the rare occasions this occurs (i.e., at most 2% of the number of the matched cases in each market). We then take the log of each instrument. The method we use is panel fixed effects instrumental variable (FEIV) where two-stage least squares are applied to within transformation data.

We provide FE and FEIV estimates in Table 2.2. To correct for correlation across products in the same province, we cluster standard errors at the province level. The coefficient on log price is significantly positive in the FE model but negative in the FEIV model, implying that the instrument variable removes a positive bias in the log(price) coefficient that exists even after the within transformation (FE model). Indeed, the regional period average prices strongly explain the log prices with high values of the F-statistics for the excluded instruments. For both specifications, the coefficients on log group shares are highly significant and less than one, suggesting that a nested demand model is appropriate.

We calculate the elasticity of substitutions within (σ) and between (γ) sub-groups using the estimated parameters from FEIV. The elasticity of substitution within sub-groups is over 2.5 times as much as between sub-groups, suggesting that modeling substitution patterns is important in market analysis. We also retrieve barcode fixed effects[12] using FEIV estimates to obtain b_{jc} for each product j purchased in region c. For singleton groups[13] that are automatically dropped in the estimations, we obtain the measurement of quality in the same way as for non-singleton groups using the estimated coefficients that are homogeneous across all barcode products. Examining the overall and within sub-group distributions of values for b_{jc}, however, we find that quality values of some singleton groups are too high. Given that these products were purchased only once in some markets during three years

[12]The individual fixed effects are inconsistent but unbiased in a large sample size with a fixed time.

[13]Under the definition of panel individuals in our model, singleton group means that product j in region c was consumed only once across all periods in that region. We do not drop the singleton groups when we construct price indexes in order to keep market variety observed in the data and avoid potential biases in price indexes.

of data collection, b_{jc} for the singleton groups is top coded at the 99th percentile of the quality measurement of each corresponding sub-group of non-alcoholic beverages to prevent the outliers from distorting price indexes. We then average the values for prices for each province over 13 quad-weeks periods of each year to build the panel price indexes at the yearly level.

2.4.2 Sample Size Bias Correction

It is possible for sample size biases to rise when more households in large cities are sampled so that more varieties of products are likely to be observed. Despite evidence from previous studies (Handbury and Weinstein, 2015; Feenstra et al., 2017) that larger cities have broader sets of product variety (using Nielsen barcode data for the U.S. and China), it is important to correct the sample size bias, even if it is minimal. This is important in our analysis because our price indexes that account for product variety can be sensitive to differences in the number of unique products observed between two chosen markets.

To correct the sample size bias, we exploit the relationship between product variety and the number of sample households in the data. We therefore specify the following regression:

$$\ln(V_m) = \beta_0 + \beta_1 \ln\left(\text{Sample}\right)_m + \theta_y + \epsilon_m \tag{2.17}$$

In equation (2.17), V_m is the number of unique UPC products observed in market m ('region-year') and $(\text{Sample})_m$ represents the number of sample households for market m. We include year fixed effects (θ_y) and cluster the standard errors at the province level.

Using the coefficients in equation (2.17) including the coefficient on $\ln(\text{Sample})_m$ (i.e., an elasticity of variety with respect to the sample size, or β_1), we can obtain an estimated count of different UPCs that is expected to be purchased by a certain number of sample households for each market. We then use the estimated UPC count for each market to limit market-specific product variety sets when markets have more observed varieties than the estimated ones. Yet we need evidence to ensure that our data from the households sample represents actual market situations, therefore providing a credible value of the elasticity of variety.

Therefore, we examine if there is supporting evidence for the use of β_1 by estimating the *total* number of varieties in each market m and comparing two elasticities of variety with respect to province size (exclusively, urban areas),[14] one from observed and the other from estimated varieties. Handbury and Weinstein (2015) introduces two methods to estimate the total number of different products based on the sample counts: 1) a parametric approach and 2) a structure approach. We use the latter, which is less intuitive than the former, because the simpler approach is only available for household level purchase data.[15] In consideration of the main purpose of this subsection, we relegate details of the structural approach to Appendix B.

We refer to the estimated total number of unique barcode products available in each market as "generalized negative exponential (GNE) estimates" because we assume that the function defining the number of different products given sample size follows the negative exponential functional form. We test if the relationship between the number of varieties available in the sample and province size (measured by the yearly urban population) is analogous to the relationship when we replace the sample varieties with the GNE estimates.

Column 1 and 2 of Table 2.3 present the results from regressing the log sample and estimated (GNE) varieties on the log of the urban population and the log of per capita income, respectively. We find a positive relationship between product variety and urban population in both regressions, echoing the results of Handbury and Weinstein (2015). We also find that the elasticity of variety with respect to urban population is larger by 0.055 % point using the sample varieties.[16] This difference is similar to what Handbury and Weinstein

[14]Note that Kantar Worldpanel collected purchase data for *urban* areas of each province.

[15]The structure approach is based on Mao et al., (2004, 2005) that suggest a methodology to estimate the number of different species in a general area based on the number of species observed in sample locations. Handbury and Weinstein (2015) shows that the parametric and structure methods provide similar results.

[16]This is because the GNE estimator corrects for the association between sample size and population in the Kantar data, as it did for Handbury and Weinstein (2015)'s Nielsen data.

(2015) found (i.e., 0.03 % point) with Nielsen U.S. scanner data, suggesting that we are able to use β_1 in equation (2.17) using our data for variety adjustment.

Despite the relatively small difference in elasticities of variety from the sample and the GNE estimates, we decide on a conservative approach by adjusting household sample size for some markets in which the probability of a household being chosen as a sample (i.e., a sampling rate) is high. Accordingly, we calculate the ratio between the number of sample households to the number of total households in each province (exclusively, urban areas) using the information Kantar provides. We then replace the sampling rates of three years of Beijing and Tianjin markets, and Shanghai-2011 with the sampling rate of the base market to obtain new sample size because these major markets with broad varieties also have high sampling rates.

By defining the adjusted household sample size as $(\text{Sample})_m$, we estimate equation (2.17) and report the result in Table 2.3, column 3. We find a positive relationship between the sample size and the number of observed unique UPC products. Then, we estimate the number of products purchased by market-specific sample size[17] using the coefficients on the variables in the model including β_1. We replace the initial number of varieties for a market with its estimated number of products when the market's original number of products is greater than the estimated variety. We have 40 markets that are classified as such.[18]

After adjusting product variety according to the relationship between the sample size and UPC count, for each market we allocate the overall number of unique products to the number of different products in each sub-group, using the sub-group shares within the initial overall UPC count. Suppose the overall unique UPC count (before the variety adjustment) in market A is 200 with 50 CSD (25%), 100 Juice (50%), 50 Tea(25%). Then, within a new overall UPC count of 100 (after the variety adjustment), the numbers of unique CSD, Juice, and, Tea are 25 (25%), 50 (50%), and 25 (25%), respectively. Then, from the original sample

[17] Note that we adjust the households sample size for Beijing, Tianjin, and Shanghai-2011 markets.

[18] These markets do not include the base market, Gui Zhou-2011. We expect it to be the case because the number of varieties and the household sample size of the base market are the smallest among all other markets.

of each market, we randomly draw samples of CSD, JUICE, and TEA for 100 times with the adjusted size of the varieties for corresponding sub-groups, from which we obtain 100 market specific NL-NCES price indexes and decompositions. We have final values of the price indexes and decompositions by taking the average of the 100 results.

2.5 Results

2.5.1 Results of Price Index and Decompositions

We construct NL-NCES price indexes by substituting quality parameters (b_{jm}[19]) along with yearly-level prices (P_{jm}) for the base and comparison markets into equation (2.4). Overall, we obtain 69 panel price indexes for 69 markets (23 provinces $\times$ 3 years). As explained in Section 2.3, the price indexes are comparable one another because they share the same market, Gui Zhou-2011, as the base. The results are presented in Table 2.4, column 1. Gui Zhou-2011 is market number 22 reported as a standardized value of 1.

One would have to multiply the prices of non-alcoholic beverage goods available in the base market by the index number in order to make households as well off as they were with the choice set of the corresponding comparison market. For instance, 0.930 index for Hubei-2011 market (Market ID number 34) means that the non-alcoholic beverage prices of the base market would have to decrease by 7% in order to give the same levels of welfare of market 34 to the households in the base market. If we want to compare any two markets other than the base market, we simply divide one market by the other. For example, if the comparison market is Jiang Su-2013 (Market ID number 42) and the base market is Jiang Su-2012 (Market ID number 41), the price index will be 1.009 (π_{42}/π_{41}), meaning that the prices of Jiang Su-2012 would have to increase by 0.9% in order to make consumers as well of as they were in Jiang Su-2013.

[19]The quality measures are time-invariant, therefore, b_{jm} equals to b_{jc} within same province c across years.

As shown in Section 2.3, the NL-NCES price decompositions show sources of welfare gaps between two markets *beyond* the simple differences in products prices. We report the effects of price, quality, and, variety that underlie the price index in Table 2.4, column 2-4, respectively. We interpret each component in the same way we do the aggregated price indexes. For instance, the variety effect for Gui Zhou-2012 (0.960) indicates that the prices of the base market (Gui Zhou-2011) would have to decrease by 4% to make up for broader variety exhibited in the same province after a year. Similarly, the quality effect (1.036) of Gui Zhou-2012 suggests that the prices of the base market would have to increase by 3.6%, implying that product variety in Gui Zhou increases from 2011 to 2012 while product quality decreases by almost the same magnitude during this period. The sum of the minus and plus values of each decomposition component relative to one approximates to the difference between corresponding NL-NCES price index values and one.

We first focus on price decompositions results and turn our attention to NL-NCES indexes results in comparison to conventional price indexes in the next subsection. We highlight two main findings from the decompositions results. First, most of between-market welfare differentials are from product quality and variety, not from prices. Comparing prices of non-alcoholic beverage products between markets, which are not compounded with variety, we find that the products' prices are similar across all markets. Second, relative to the base market, we see examples where the effects of variety and quality on cost-of-living indexes are opposite. For instance, Shaan Xi-2012 has a similar cost-of-living as the base market but this is a result of the market's broader variety but lower product quality (given that price level between two markets is similar). Specifically, the quality effect for Shaan Xi-2012 suggests that the prices of the base market would have to *increase* by 34% to be penalized for higher quality than Shaan Xi-2012. On the other hand, the variety effect for Shaan Xi-2012 indicates that the prices of the base market would have to *decrease* by 23% to make up for broader variety of Shaan Xi-2012.

This finding implies that a broader set of varieties is not necessarily correlated with higher quality. In the context of the Chinese non-alcoholic beverage market, the opposite effects of variety and quality would be the case when many local brands exist but demonstrate lower quality compared to national/international brands. Therefore, depending on the number and quality of local non-alcoholic beverage brands in each market, we can expect that the effects of quality and variety on cost-of-living indexes may cancel each other out for some markets. For instance, while the base market has 350 different juice products, only five UPCs of which are from Gui Zhou local brands, Hu Bei-2013 market has on average 15 times as many local brands as the base market among 350 randomly drown UPCs out of 733 juice products. On the contrary, the two effects for the large cities have the same directions in that consumers in those cities enjoy broader varieties with higher quality than the base market. The positive relationship between variety and quality effects may be attributed to market openness of the large cities for which diverse products from international brands with high quality can be readily imported relative to the base market.

Overall, our results highlight the important roles of variety and quality effects in determining cost-of-living indexes when consumers are assumed to gain utility from broad variety and high quality of products. This then leads to a subsequent question: how important is it to capture variety and quality in cost-of-living indexes? We address the product heterogeneity bias in the next subsection.

2.5.2 Product Heterogeneity Bias

We examine the existence of a product heterogeneity bias and quantify the bias, if it exists, by comparing our NL-NCES indexes with two representative superlative indexes (Fisher ideal and Törnqvist). To construct superlative indexes, we identify 'identical' barcode products between the base market (Gui Zhou-2011) and chosen comparison markets and use market specific price and quantity information of the products in the common product bundles.[20]

[20] These indexes are constructed by averaging over 100 results from 100 randomly drawn samples that we used for the NL-NCES indexes (see, subsection 2.4.2).

In other words, both superlative price indexes do not account for variety and quality of products. As all 69 price indexes share the same base market, Fisher ideal and Törnqvist indexes also satisfy the transitivity. The results are presented in Table 2.5.

To facilitate the price index comparison, we illustrate the NL-NCES and Fisher price indexes (given that Törnqvist index values are similar to Fisher values) relative to Gui Zhou-2011 presented in Table 2.5, by plotting these in maps side by side by years shown in Figure 2.1 to 2.3.

We find that the failure of considering product variety and quality significantly misleads welfare differences between the base and a comparison market in a way that conventional price indexes largely underestimate the differences. The NL-NCES index values widely range from 0.45 to 1.10 while the conventional index values are analogous across markets between 0.90 and 1.03. We also find that the underestimation of welfare differences is relatively more prominent in the spatial dimension than in the temporal dimension. The large spatial price variations shown in NL-NCES indexes are comparable to those in product variety across provinces in China.

Moreover, the heterogeneity bias is significant when the comparison market is a large city with advanced economic development such as Beijing, Shanghai, and Tianjin. For instance, the NL-NCES index indicates that the prices of non-alcoholic beverages in Gui Zhou-2011 would have to decrease by 47.2% to compensate welfare shortage compared to Beijing-2011 when the magnitude of the price decrease is only 5.8% in the Fisher index. This large welfare gap in NL-NCES indexes may not be surprising given that economic development in Beijing is much more advanced than that in Gui Zhou.[21] In this regard, our findings using the non-alcoholic beverages data provide suggestive evidence for large spatial economic inequality in China.

[21] The per-capita GDPs of Beijing from 2011 to 2013 are over five times as much as that of Gui Zhou-2011. Also, the Chinese city tier system, often referred to by various media publications, classifies Beijing into the top tier and the capital city of Gui Zhou (Guiyang) into the bottom tier out of five.

We can see how the heterogeneity bias differs by province more easily in Figure 2.4 where NL-NCES indexes are plotted against Fisher indexes with the size of the markers representing provincial per-capita GDP as a proxy for economic development levels. Comparing index values between superlative and NL-NCES for provinces with major cities, we find that substantial downward biases in measuring between-market welfare gaps in superlative cost-of-living indexes amount up to 94.5%.

For some Chinese provinces with medium and small size cities, we find that the index values between the superlative indexes and ours are similar but that the product heterogeneity bias could potentially distort price indexes. This argument is based on the fact that the similar price index numbers often arise from the trade-off between the variety and quality effects (e.g., Hei Long Jiang and Hu Bei markets).

Overall, the absence of variety and quality effects in cost-of-living indexes is likely to cause underestimation of welfare gaps between comparison markets and the base market with the smallest variety, and the downward bias becomes significant among cost-of-living indexes for large cities.

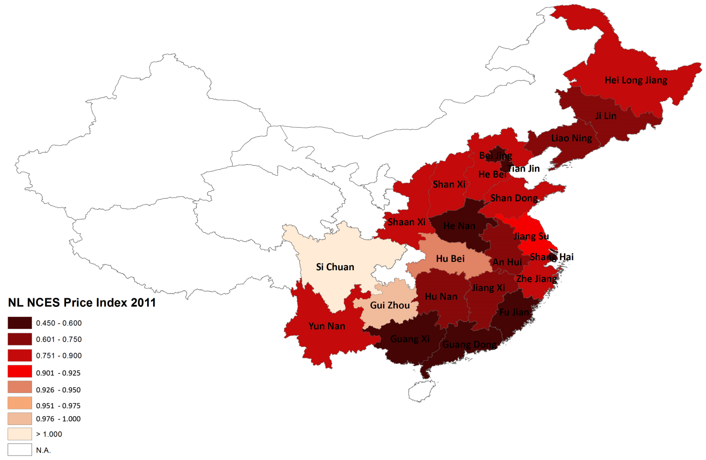
2.6 Conclusions

Conventional matched-basket price indexes measure either temporal or spatial cost-of-living variation but not both. Moreover, these indexes only partially account for product heterogeneity with respect to both quality and variety. A handful of studies show that product varieties varies across regions and over time, implying that ignoring product varieties could bias the true cost-of-living indexes, especially in large countries where product heterogeneity is likely to be substantial. Despite its importance, knowledge of the effects of variety *and* quality of products in welfare analysis is still scant, thus requring more rigorous investigation.

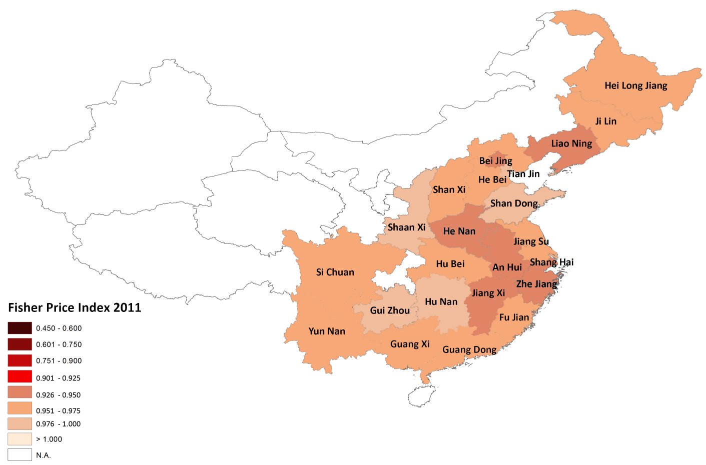
To the best of our knowledge, this paper is the first attempt to investigate the magnitude of a bias when we ignore the effects of product variety *and* quality on cost-of-living indexes in the context of a large developing country. We build the panel price indexes using a large-scale retail scanner data of non-alcoholic beverages for 23 provinces over three years in China. Setting a market with the smallest varieties as the base market, our panel price indexes are comparable for any two chosen markets. Decomposing the price indexes into three components, we disentangle the effects of price and quality from the effect of variety with a novel sampling method.

We find that quality and variety are the primary components to contribute to differentials in between-market cost-of-living while the price level of non-alcoholic beverage products are analogous across markets. Comparing our indexes with two superlative price indexes in which product quality and variety are not considered, we confirm that ignoring quality and variety effects is likely to underestimate cost-of-living and by extension, the welfare gaps between two markets. This downward bias is largest among price indexes of Chinese major cities such as Beijing, Shanghai, and, Tianjin.

From a policy perspective, our results have a potentially important implication. If cost-of-living measurement in the absence of the variety and quality effects provides misleading pictures of regional economic gaps, it is possible that region-specific policies with the aim of reducing these gaps would be less effective than expected when the target region is incorrectly identified. Given that spatial inequality is a common social problem both for developed and developing countries such as the United States, China and India, advancements in spatial price indexes that address the two dimensions of product heterogeneity would increase the effectiveness of the policies because policymakers have better information about which province is most underdeveloped with a source of the high welfare gap.



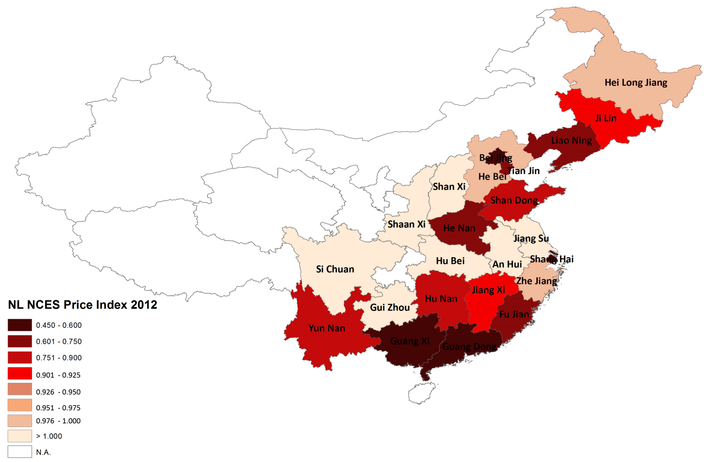
(a) NL-NCES Price Index



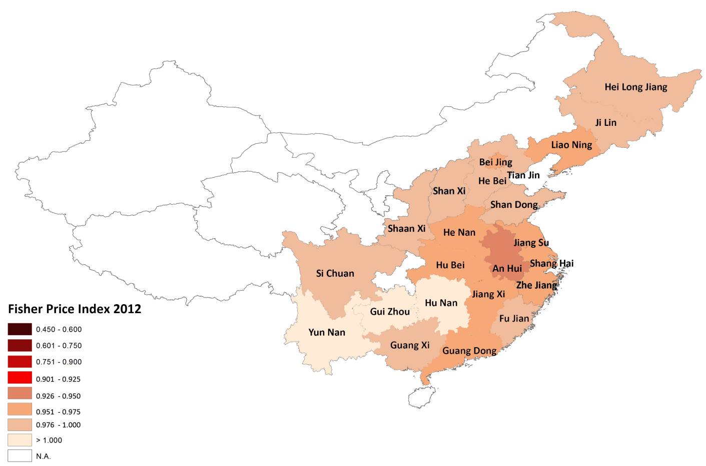
(b) Fisher Ideal Price Index

Figure 2.1: 2011 Cost-of-living in China: Non-alcoholic Beverages

Notes: This figure illustrates the 2011 NL-NCES and Fisher Ideal price indexes for non-alcoholic beverages relative to the base market (Gui Zhou-2011) across 23 Chinese provinces. Note that we construct the indexes based on the urban population's purchases data. The price index number for each province can be found in Table. 5.



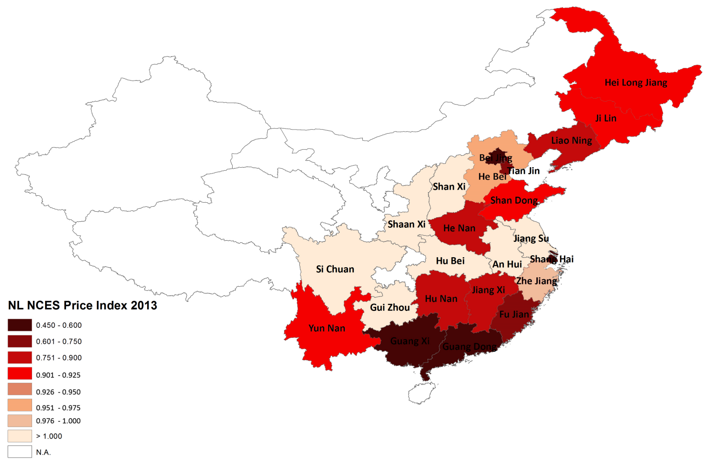
(a) NL-NCES Price Index



(b) Fisher Ideal Price Index

Figure 2.2: 2012 Cost-of-living in China: Non-alcoholic Beverages

Notes: This figure illustrates the 2012 NL-NCES and Fisher Ideal price indexes for non-alcoholic beverages relative to the base market (Gui Zhou-2011) across 23 Chinese provinces. Note that we construct the indexes based on the urban population's purchases data. The price index number for each province can be found in Table. 5.



(a) NL-NCES Price Index

(b) Fisher Ideal Price Index

Figure 2.3: 2013 Cost-of-living in China: Non-alcoholic Beverages

Notes: This figure illustrates the 2013 NL-NCES and Fisher Ideal price indexes for non-alcoholic beverages relative to the base market (Gui Zhou-2011) across 23 Chinese provinces. Note that we construct the indexes based on the urban population's purchases data. The price index number for each province can be found in Table. 5.

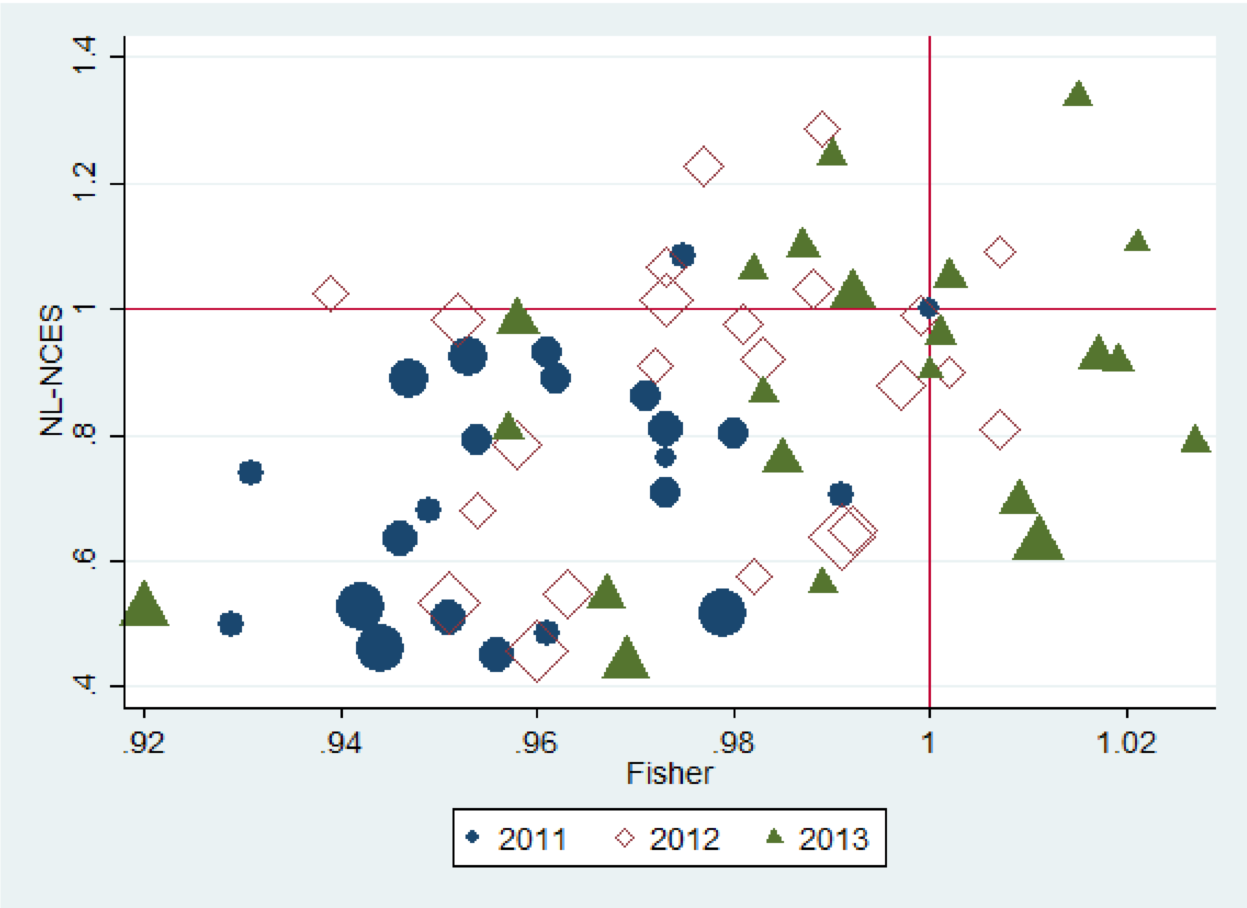


Figure 2.4: Heterogeneity Bias by Province and Year

Notes: This figure illustrates how the heterogeneity bias differs by province and year by plotting NL-NCES indexes against Fisher indexes with the size of the markers representing provincial per-capita GDP as a proxy for economic development levels.

Table 2.1: Summary Statistics for Kantar Worldpanel Data

(Total sample size: 333,877)

UPC count

	Min	Med	Max
UPC per Province	1184	2277	3092
UPC per Period	3180	4150	4753
UPC per Province & Period	116	392	882
	CSD	JUICE	TEA
UPC per Sub-group	3921	10585	2114

Notes: This table presents the summary statistics for the sample we use to estimate demand. The barcode-level non-alcoholic beverages purchasing dataset were collected by Kantar from a sample of 39,000 households in 23 Chinese provinces at the quad-week period level during 2011 to 2013. The non-alcoholic beverage products are categorized into seven sub-groups: Carbonated soft drinks (CSD), Juice, Ready-to-drink (RTD) tea, Functional drinks, Packaged water, RTD coffee, and Soybean milk. We use three subcategories (i.e., CSD, JUICE and TEA) among seven that allow us to have one common base market in terms of spatial and time dimensions. The number of unique UPC varies widely across provinces and periods providing sufficient variations to identify demand parameters.

Table 2.2: Modified Logit Demand Regression Results

Specification:	FE	FEIV
log(price)	0.144*	-0.146***
	(0.075)	(0.047)
log(group share)	0.921***	0.926***
	(0.016)	(0.016)
Implied Elasticities of Substitution:		
γ		1.146***
		(0.047)
σ		2.963***
		(0.894)
F-statistics for instrument		953.74
p-value		0.000
Observations	317,968	317,968

Notes: This table shows the main parameter estimates of the modified logit demand equation: one is estimated with panel fixed effects (FE) and the other is estimated with panel fixed effects instrumental variable (FEIV) methods. Singletons are automatically omitted in the fixed effect analysis. The coefficients on log price is significantly positive in the FE model but negative in the FEIV model, implying that the instrument variable removes a positive bias in the log (price) coefficient. The F-statistics for the excluded instruments are high. Based on the parameters of the FEIV model, we calculate the elasticity of substitutions within (σ) and between (γ) sub-groups. Standard errors in parentheses are clustered at the province level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.3: Regressions for Sample Size Correction

	$\ln(V)_m$	$\ln(GNE)_m$	$\ln(V)_m$
$\ln(\text{Sample})_m$			0.390^{***}
			(0.050)
$\ln(\text{Urban pop})_m$	0.243^{***}	0.188^{***}	
	(0.058)	(0.057)	
$\ln(\text{Per Capita Income})_m$	0.630^{***}	0.437^{***}	
	(0.155)	(0.145)	
Year fixed effects	Y	Y	Y
Observations	69	69	69
R-squared	0.501	0.370	0.688

Notes: This table includes regressions used for correcting the sample size bias. The first two regressions test if the data from the households sample represents actual market situation. Specifically, we first estimate the elasticity of variety with respect to province size (exclusively, urban areas) using the UPC count from the sample (V_m). We then estimate the total number of product varieties (GNE_m) in each market using the sample UPC count and purchase frequency to compare the elasticity of variety from V_m with that from GNE_m (column 1 and column 2). Once we confirm a little difference between the two elasticities, we exploit the relationship between product variety and the number of sample households (Sample_m) in the data, which is shown in column 3. We use coefficients on parameters (including year fixed effects) to obtain an estimated UPC count for each market which is used to adjust market-specific sample varieties. Standard errors in parentheses are clustered at the province level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.4: NL-NCES Price Indexes and Price Index Decompositions

		Market	**NL-NCES**	Price Index Decompositions		
Province	Year	ID	**Price Index**	**Price**	**Quality**	**Variety**
Bei Jing	2011	1	0.528	0.927	0.736	0.703
Bei Jing	2012	2	0.533	0.963	0.735	0.690
Bei Jing	2013	3	0.523	1.041	0.735	0.671
Shang Hai	2011	4	0.461	0.886	0.756	0.653
Shang Hai	2012	5	0.458	0.936	0.739	0.642
Shang Hai	2013	6	0.437	0.973	0.701	0.620
Tian Jin	2011	7	0.518	0.998	0.615	0.813
Tian Jin	2012	8	0.639	1.007	0.728	0.804
Tian Jin	2013	9	0.627	1.024	0.724	0.781
An Hui	2011	10	0.740	0.932	1.235	0.721
An Hui	2012	11	1.027	0.993	1.433	0.717
An Hui	2013	12	1.064	1.003	1.500	0.695
Fu Jian	2011	13	0.511	0.941	0.701	0.761
Fu Jian	2012	14	0.649	0.990	0.813	0.738
Fu Jian	2013	15	0.694	1.024	0.888	0.703
Guang Dong	2011	16	0.450	1.012	0.859	0.536
Guang Dong	2012	17	0.547	1.078	0.917	0.529
Guang Dong	2013	18	0.545	1.110	0.927	0.522
Guang Xi	2011	19	0.486	0.982	0.617	0.734
Guang Xi	2012	20	0.576	0.975	0.674	0.767
Guang Xi	2013	21	0.565	1.009	0.669	0.740
Gui Zhou	2011	22	1	1	1	1
Gui Zhou	2012	23	1.091	1.084	1.036	0.960
Gui Zhou	2013	24	1.105	1.110	1.072	0.891
He Bei	2011	25	0.792	1.038	0.945	0.742
He Bei	2012	26	0.977	1.040	1.036	0.729
He Bei	2013	27	0.961	1.092	1.043	0.707
Hei Long Jiang	2011	28	0.862	0.960	1.097	0.767
Hei Long Jiang	2012	29	0.992	0.973	1.169	0.756
Hei Long Jiang	2013	30	0.916	1.018	1.160	0.723
He Nan	2011	31	0.501	0.982	0.767	0.673
He Nan	2012	32	0.679	1.036	0.894	0.660
He Nan	2013	33	0.810	1.067	1.046	0.638
Hu Bei	2011	34	0.930	0.912	1.301	0.736
Hu Bei	2012	35	1.068	0.938	1.454	0.724
Hu Bei	2013	36	1.100	0.993	1.487	0.702
Hu Nan	2011	37	0.706	0.966	0.790	0.866
Hu Nan	2012	38	0.810	1.031	0.874	0.814

Continued on next page

Table 2.4: NL-NCES Price Indexes and Price Index Decompositions

		Market	**NL-NCES**	Price Index Decompositions		
Province	Year	ID	**Price Index**	**Price**	**Quality**	**Variety**
Hu Nan	2013	39	0.790	1.051	0.907	0.801
Jiang Su	2011	40	0.923	0.909	1.599	0.612
Jiang Su	2012	41	1.015	0.957	1.698	0.613
Jiang Su	2013	42	1.024	0.984	1.672	0.593
Jiang Xi	2011	43	0.679	0.929	0.984	0.784
Jiang Xi	2012	44	0.909	0.997	1.097	0.759
Jiang Xi	2013	45	0.867	1.042	1.090	0.747
Ji Lin	2011	46	0.707	0.945	0.888	0.829
Ji Lin	2012	47	0.922	1.012	1.020	0.815
Ji Lin	2013	48	0.914	0.975	1.104	0.775
Liao Ning	2011	49	0.635	0.960	0.977	0.655
Liao Ning	2012	50	0.784	1.006	1.090	0.657
Liao Ning	2013	51	0.760	1.048	1.115	0.628
Shaan Xi	2011	52	0.802	0.917	1.207	0.777
Shaan Xi	2012	53	1.032	0.995	1.341	0.785
Shaan Xi	2013	54	1.050	1.023	1.388	0.751
Shan Dong	2011	55	0.811	1.005	1.099	0.645
Shan Dong	2012	56	0.880	1.064	1.126	0.637
Shan Dong	2013	57	0.925	1.102	1.150	0.614
Shan Xi	2011	58	0.889	0.987	1.090	0.818
Shan Xi	2012	59	1.226	1.035	1.292	0.805
Shan Xi	2013	60	1.246	1.051	1.359	0.791
Si Chuan	2011	61	1.083	0.939	1.521	0.728
Si Chuan	2012	62	1.287	1.021	1.678	0.700
Si Chuan	2013	63	1.337	1.053	1.777	0.681
Yun Nan	2011	64	0.763	0.999	0.971	0.775
Yun Nan	2012	65	0.900	1.051	1.061	0.760
Yun Nan	2013	66	0.903	1.090	1.045	0.750
Zhe Jiang	2011	67	0.890	0.923	1.424	0.646
Zhe Jiang	2012	68	0.985	0.982	1.452	0.646
Zhe Jiang	2013	69	0.981	1.036	1.415	0.631

Notes: This table represents NL-NCES price indexes for 69 markets (23 provinces $\times$ 3 years) in column 4 and NL-NCES price decompositions in column 5-7. The price decomposition shows the respective effect of price, quality, and variety on welfare gaps between two markets (a comparison and the base market).

Table 2.5: NL-NCES vs. Superlative Price Indexes

		Market	Superlative Price Index		**NL-NCES**
Province	Year	ID	**Fisher**	**Törnqvist**	**Price Index**
Bei Jing	2011	1	0.942	0.940	0.528
Bei Jing	2012	2	0.951	0.951	0.533
Bei Jing	2013	3	0.920	0.919	0.523
Shang Hai	2011	4	0.944	0.942	0.461
Shang Hai	2012	5	0.960	0.959	0.458
Shang Hai	2013	6	0.969	0.969	0.437
Tian Jin	2011	7	0.979	0.980	0.518
Tian Jin	2012	8	0.991	0.993	0.639
Tian Jin	2013	9	1.011	1.012	0.627
An Hui	2011	10	0.931	0.932	0.740
An Hui	2012	11	0.939	0.938	1.027
An Hui	2013	12	0.982	0.982	1.064
Fu Jian	2011	13	0.951	0.950	0.511
Fu Jian	2012	14	0.992	0.990	0.649
Fu Jian	2013	15	1.009	1.008	0.694
Guang Dong	2011	16	0.956	0.955	0.450
Guang Dong	2012	17	0.963	0.962	0.547
Guang Dong	2013	18	0.967	0.966	0.545
Guang Xi	2011	19	0.961	0.960	0.486
Guang Xi	2012	20	0.982	0.980	0.576
Guang Xi	2013	21	0.989	0.988	0.565
Gui Zhou	2011	22	1	1	1
Gui Zhou	2012	23	1.007	1.007	1.091
Gui Zhou	2013	24	1.021	1.021	1.105
He Bei	2011	25	0.954	0.955	0.792
He Bei	2012	26	0.981	0.983	0.977
He Bei	2013	27	1.001	1.002	0.961
Hei Long Jiang	2011	28	0.971	0.974	0.862
Hei Long Jiang	2012	29	0.999	1	0.992
Hei Long Jiang	2013	30	1.019	1.020	0.916
He Nan	2011	31	0.929	0.929	0.501
He Nan	2012	32	0.954	0.954	0.679
He Nan	2013	33	0.957	0.958	0.810
Hu Bei	2011	34	0.961	0.961	0.930
Hu Bei	2012	35	0.973	0.972	1.068
Hu Bei	2013	36	0.987	0.987	1.100
Hu Nan	2011	37	0.991	0.989	0.706
Hu Nan	2012	38	1.007	1.007	0.810

Continued on next page

Table 2.5: NL-NCES vs. Superlative Price Indexes

Province	Year	Market ID	Superlative Price Index Fisher	Superlative Price Index Törnqvist	**NL-NCES Price Index**
Hu Nan	2013	39	1.027	1.027	0.790
Jiang Su	2011	40	0.953	0.953	0.923
Jiang Su	2012	41	0.973	0.973	1.015
Jiang Su	2013	42	0.992	0.992	1.024
Jiang Xi	2011	43	0.949	0.949	0.679
Jiang Xi	2012	44	0.972	0.971	0.909
Jiang Xi	2013	45	0.983	0.984	0.867
Ji Lin	2011	46	0.973	0.976	0.707
Ji Lin	2012	47	0.983	0.985	0.922
Ji Lin	2013	48	1.002	1.004	0.914
Liao Ning	2011	49	0.946	0.947	0.635
Liao Ning	2012	50	0.958	0.959	0.784
Liao Ning	2013	51	0.985	0.987	0.760
Shaan Xi	2011	52	0.980	0.980	0.802
Shaan Xi	2012	53	0.988	0.987	1.032
Shaan Xi	2013	54	1.002	1.002	1.050
Shan Dong	2011	55	0.973	0.975	0.811
Shan Dong	2012	56	0.997	1.002	0.880
Shan Dong	2013	57	1.017	1.019	0.925
Shan Xi	2011	58	0.962	0.962	0.889
Shan Xi	2012	59	0.977	0.978	1.226
Shan Xi	2013	60	0.990	0.992	1.246
Si Chuan	2011	61	0.975	0.975	1.083
Si Chuan	2012	62	0.989	0.989	1.287
Si Chuan	2013	63	1.015	1.015	1.337
Yun Nan	2011	64	0.973	0.974	0.763
Yun Nan	2012	65	1.002	1.001	0.900
Yun Nan	2013	66	1	1	0.903
Zhe Jiang	2011	67	0.947	0.945	0.890
Zhe Jiang	2012	68	0.952	0.951	0.985
Zhe Jiang	2013	69	0.958	0.957	0.981

Notes: This table shows two superlative price indexes (Fisher and Törnqvist) and the NL-NCES price indexes that account for product quality and variety. The two superlative price indexes do not account for product heterogeneity. By comparing price indexes between superlative and NL-NCES, therefore, we quantify the product heterogeneity bias.

Chapter 3

New estimates of consumer demand for foods and nutrients across households in Tanzania

3.1 Introduction

Malnutrition continues to be a great challenge in developing countries, causing adverse long-term effects on health and the development of human potential at the individual level, and social and economic development and growth at the national level. Although the global consensus involving the importance of combating the problem is well established, nutrition interventions are often promoted based on limited understanding of how households' food choices shape nutrient status. For instance, much of the existing literature on diet quality focuses on increased consumption of a few food items on nutrition improvement (e.g., Ruel et al. 2005; Weinberger and Swai 2006; Ochieng et al. 2018). This perspective may not adequately capture true dynamics of households overall nutrient intake by missing complex substitutive and complementary relations among foods. Moreover, most consumers do not demand foods primarily based on nutrient content, but rather based on their personal preferences for certain food and non-food products, their available income, and the prices reflected in the marketplace. Because food products are the primary delivery mechanism for dietary macro and micro nutrients, it is critical to understand how consumer preferences on a variety of foods mediate the pathways between the economic shocks and households' diets.

In this paper, we use a holistic approach to advance our knowledge of household' food consumption and nutrient intake sensitivity to changes in food prices and total expenditure. Specifically, we estimate a large food demand system comprised of 19 food categories and a numéraire good using a two-way Exact Affine Stone Index (EASI) model developed by

Lewbel and Pendakur (2009). The EASI model shares the advantages of the popular Almost Ideal Demand (AID) model (Deaton and Muellbauer, 1980) such as an approximation to a linearized model accounting for censoring, and also provides additional flexibility that is not provided by the AID model. With the EASI model, the empirical data defines the shape of the Engel curves and can reveal flexible price effects conditional on total expenditure.

Using a nationally representative panel household survey in Tanzania, we first estimate a cross-sectional EASI model and obtain food consumption elasticities with respect to food prices and total expenditure. We then extend the model to a panel structure with a correlated random effects specification (Meyerhoefer et al., 2005) to account for unobserved time-invariant regional characteristics. Estimating food demand systems with these two models creates a unique opportunity to compare resulting elasticities between the models and evaluate how controlling the unobserved heterogeneity alters the estimates. Following Huang (1996) and Huang and Lin (2000), we also derive nutrient intake elasticities for energy and seven key nutrients (i.e., protein, fat, sugar, iron, zinc, vitamin A, and total folate) using the estimated price and total expenditure elasticities from the panel model.

We do so not only for the aggregated sample but also for four different per-capita total expenditure groups. The quartile level analysis provides better knowledge of how food and nutrient consumption demand of relatively poorer households respond differently to changes in food prices and total expenditure when compared to wealthier households. Hence, findings from our study are particularly informative for policy makers because it points to how nutrition improvement policies may generate large positive effects with minimal side-effects only when they are well designed for target households based on a good knowledge about the potential beneficiaries.

Comparing estimated elasticities from the two models, we first find that ignoring the unobserved heterogeneity is likely to cause an upward bias to own-price elasticities in absolute sense and over-rejection of the null cross-price effects. Considering the fact that households reveal strong preferences on particular foods that are easily accessible and grown in their regions (Atkin, 2013) and hence the price of the foods remains relatively low, these results reinforce the importance of accounting for price variations attributed by regional tastes differences. Notably, our results do not demonstrate a distinction between the models in the total expenditure elasticities results.

According to the elasticities estimated from the panel model, we find that the absolute magnitudes of own-price elasticities generally fall in rising total expenditure, with the exception of food groups on which wealthier households have either low preferences (e.g., cassava) or enough consumption (e.g., vegetables, dairy, coffee, tea, and cocoa, other food). We also find that poorer households have more food substitutes compared to more wealthier households. This is consistent with prior knowledge suggesting that limited-resource households are more willing to substitute to reduce the overall cost of food. This finding also demonstrates the importance of deploying a systematic approach in food consumption analysis, and by extension, nutritional analysis.

Total expenditure elasticities are mostly significant across all quartiles and have greater magnitudes than price elasticities in terms of the amount of changes in food consumption quantity. This finding suggests that income support policies would be more effective than price policies in modifying the level of consumption for overall food groups. Due to tight linkages between food consumption and nutrient intake, we confirm that this result also applies to the nutrition consumption. The calculated nutrient price elasticities are predominantly small and insignificant but nutrient total expenditure elasticities are large and highly significant, placing an emphasis on income growth for both food security and nutrition improvement.

We further explore how nutritional impacts of food prices and total expenditure differ by per-capita total expenditure. Although nutritional effects of food prices are mostly small and insignificant, we find that poor households could benefit from price-related policies for staples such as maize and pulses for overall nutrients intake. Total expenditure effects on nutrient intake is decreasing in response to declining per-capita total expenditure. Moreover, for the bottom 25% households, we note the severe imbalance in the nutritional effect of additional total expenditure, especially for micronutrients. The poorest households disproportionally increase consumption of iron, zinc, vitamin A, and total folate while over-proportionally increase consumption of fat and sugar.

The unbalanced nutritional effects of total expenditure among the poorest households confirm the notion that individuals select foods based on preference not perceived nutritional benefit. This finding implies that the poor have strong preferences for relatively unhealthy foods and are likely to spend more on these foods when they have additional income. Depending on the underlying mechanism for this phenomenon, we draw the following different but equally important policy implications for balanced nutrition improvement. If a lack of understanding about nutritional benefits is the primary driver, then providing nutrition education, particularly focusing on benefits of micronutrients, would promote healthier food choices of the poor. If limited access to micronutrient-dense foods is the primary driver, then food transfer programs for higher micronutrients consumption may be more effective and suitable. The direct food transfer programs, if implemented over a long period, will also promote the poor households to form habit of consuming nutrient-rich foods in the long run.

This study addresses an important gap in the literature on food consumption and nutritional analysis. First, this study contributes to an under-developed area of the literature by modeling a comprehensive food demand system in the context of a Sub-Saharan African country. Although large scale food demand system models are commonly used in studies of developed countries, their use in studies of developing countries, especially Sub-Saharan Africa, has been much more limited. Even those few papers that estimate demand systems of Sub-Saharan countries, their results suffer from potential biases due to issues regarding either conditional demand model estimation (e.g., conditional on food expenditure) or price endogeneity (Weliwita et al., 2002; Abdulai and Aubert, 2004; Lazaro et al., 2017). Our study overcomes these potential econometric problems by estimating an incomplete demand system with a composite numéraire good and constructing Fisher Ideal indexes as well as their instruments for price.

Second, to our best knowledge, our study is the first to estimate a large and flexible food demand system in a panel structure for a developing country. Among Tanzanian studies, Weliwita et al. (2002) estimate the linearized AID system using the Tanzania 1991/92 household budget survey data and Abdulai and Aubert (2004) also estimate price and expenditure elasticities for Tanzanian using the Quadratic AID (Banks et al., 1997), but both studies rely on cross-sectional settings. Other food demand studies such as Agbola (2003), Bopape and Myers (2007) (South Africa), and Ecker and Qaim (2011) (Malawi) are all based on cross-sectional model analysis. By comparing cross-sectional model results to panel ones, we provide evidence that a failure to address regional taste differences biases own-price elasticities.

The rest of the paper proceeds as follows. Section 3.2 provides an introduction to the data. Section 3.3 briefly presents the demand system model as well as its extension. Section 3.4 describes econometrics strategy to address problems in estimating the models from Section 3.3 and derivations of nutrient elasticities. Section 3.5 reports results of the food consumption and nutrient intake elasticities with respect to food price and total expenditure changes. The final section concludes.

3.2 Data Overview

We use data on household characteristics and food consumption in the Tanzania Living Standards Measurement Study (LSMS) program. The LSMS is a global integrated household survey program within the World Bank's Development Data Group, collating various social and economic variables from nationally representative households in collaboration with national statistical offices. Beyond the nationally representative sample and extensive survey questionnaires, the LSMS has the advantage of building household survey data in a panel structure. The survey was designed so that original households surveyed at the first round were targeted for re-interviewed and any adults within the original households who created new households were also tracked with different household ID.[1] This panel structure allows us to control for unobserved time-invariant household characteristics in a model, and by extension, compare cross-sectional results with panel ones. The longitudinal sample was refreshed in the fourth round due to changes in administrative boundaries and demographic information, therefore, we use the first three rounds of the survey data that span the periods of October 2008 – October 2009, October 2010 - September 2011, and October 2012 - November 2013, respectively.

The sample comprises 11,115 households in a cross-sectional structure,[2] which equates to 5,018 households in a panel structure. The sample was constructed based on the National Master Sample frame which is a list of all populated enumeration areas (EA) in the country developed from the 2002 Population and Housing Census (Basic Information Document-National Panel Survey (NPS 2008-2009), 2009). Samples are also classified into 410 EAs, 26 regions, 8 administrative zones, and 3 urban types (Dar-es-Salaam/other urban/rural). Particularly, the combination of the administrative zones and urban/rural provides 16 strata at which sampling weights for nationally representative statistics are calculated.

The household food consumption module assembles data on consumed quantities for 56 food items (excluding alcoholic beverages) using a 7-day recall. Households' total quantity consumed for each food item is then decomposed according to three different sources: market purchase, own production, and food gifts. If some of the consumed quantities are from market purchase, the survey reports expenditures for purchased foods in Tanzanian Shilling (TSH). Given the absence of disaggregated product-level prices (e.g., barcode product), we use unit values as prices that are calculated by dividing the expenditures by the food quantities from market purchase. The use of unit values introduces an econometric problem, which we discuss in Section 3.4.

[1] The split-off household is considered as an independent panel household apart from its original household.

[2] The initial total number of households is 12,199 but we remove households whose daily per capita kcal, calculated based on household composition and adult-equivalent scales, is less than 1% or greater than 95% of the distribution of daily per capita kcal in a corresponding survey round.

We aggregate 56 food items into 19 food-at-home (FAH) categories based on the typical composition of Tanzanian meals and the nutritional characteristics of foods. The 19 food categories are rice; maize; wheat and other cereals; cassava; roots, tubers and other starches; sugar; pulses; nuts and seeds; vegetables; fruit; red meat; poultry; eggs; fish and seafood; dairy; fats and oils; coffee, tea, and cocoa; soft drink and juice; and other food (salt and other spices). For consistency of variable creation and interpretation, we choose one measurement unit at which the majority food items within each food category are reported in the survey and standardize the measurement unit (either kilograms, litre, or pieces) at the food category level.

The LSMS data also provides information on households' expenditures on non-food items from which we calculate households' total expenditures and budget share on foods. We use the total expenditures to divide households into four groups according to the quartiles of the per-capita total expenditure distribution. Figure 3.1 presents households' average budget share on foods by per-capita total expenditure quartile. As expected, the budget share on foods decreases with rising per-capita total expenditure. The bottom quartile households spend 76% of their total expenditures on foods while the top quartile households only spend 46% on foods.

We further report descriptive statistics of the average budget share of the total food budget and unit value by food category and per-capita total expenditure quartile in Table 3.1. Among all food groups, maize accounts for the highest budget share except for the top quartile, suggesting that maize is the staple food for most Tanzanians. The demand for cassava and pulses sharply declines as households become richer. This drastic reduction reflects the fact that poor households heavily rely on relatively cheap grains to obtain energy whereas rich households can afford to diversify their staple foods to relatively expensive crops such as rice, wheat and other cereals. For all food categories, the average unit value increases as per-capita total expenditure rises, implying that richer households increasingly consume higher quality products. The unit prices of the traditional staple foods such as maize, cassava, roots and tubers are relatively inexpensive whereas the unit prices of animal source foods and fish and seafood are expensive. Overall, with increasing income, households diversify their diet by shifting food preferences from traditional calorie-dense foods to tasty and less calorie-dense foods.

As mentioned in the introduction, we focus on three macronutrients (protein, fat, and sugar) and four micronutrients (iron, zinc, vitamin A, and total folate) in addition to calories to better understand nutritional impacts of changes in household income and food prices. From the consumed quantities for each food item that households reported, we calculate the nutrient values by applying conversion factors from various sources[3] that account for adjustments for edible portion and nutrient losses during cooking.

3.3 Model

Basic Model

We characterize Tanzania household food preferences in an EASI demand system with 19 FAH categories and a composite numéraire good that incorporates all other consumption goods and services. In contrast to a food demand conditional on food expenditures, this incomplete demand system with a composite numéraire good allows for a substantial reduction in endogeneity biases[4], thus allowing for more direct policy-relevant implications.

We choose the EASI functional form, as opposed to the popular AID model and its variants, for three primary reasons. First, like the AID model, EASI has an approximate version that is linear in parameters conditional on real total expenditures. This allows accounting for censored demand in estimation. Without this property, it would have been extremely difficult econometrically to estimate a nonlinear demand functional form with censored data.

[3]We obtain conversion factors for each of 56 food items from the following sources: 1) Tanzania Food Composition Tables, Lukmanj et al., 2008; 2) WorldFood Dietary Assessment System International Mini-list (2nd edition), University of California at Berkeley, 2006; and 3) USDA National Nutrient Database for Standard Reference, 2016.

[4]Zhen et al. (2014) shows that a conditional demand model overpredicts the reduction in total calories in response to sugar sweetened beverage taxes by failing to capture substitutions between the beverages.

Second, the EASI model allows the Engel curves to take any shape as determined by data. This feature is especially important in the context of developing economies because household income ranges are often wide and therefore, expenditure elasticities for foods can vary by income groups (e.g., the lowest 10% vs. the highest 10% of the household income distribution). By contrast, the AID model is limited to just accommodating a demand quadratic in total real expenditures (Banks et al., 1997). Third, by including interaction terms between log prices and real total expenditure in the model, which is termed as the two-way EASI, the model allows Hicksian demand to vary with total expenditure in a utility-theoretic fashion. Hence, the EASI functional form could be consistent with extremely flexible price effects in the data. The AID models only allow Marshallian demand to vary with income through the income effect in the Slutsky equation.

The two-way approximate EASI demand system is specified as

$$w^*{}_{hi} = \mu_i + \sum_{j=1}^{J} \alpha_{ij} \log p_{hj} + \sum_{r=1}^{L} \beta_{ir} {y_h}^r + \sum_{j=1}^{J} \alpha_{ijy}(y_h \times \log p_{hj}) + \sum_{k=1}^{K} \gamma_{ik} z_{hk} + u_{hi}, \tag{3.1}$$

$(h = 1, ..., H; i = 1, ...J - 1)$

In equation (3.1), w^*_{hi} is the latent budget share on the i th FAH category for household h. As the LSMS data records households' food consumption during the short recall period, zero consumption of some food groups is unavoidable. Because of censoring, the latent share $w^*{}_{hi}$ is related to observed budget share w_{hi} according to $w_{hi} \equiv \max\{0, w^*{}_{hi}\}$. The first independent variable next to constant term μ_i (p_{hj}) is the price index for household h and category j with J being the number of demand categories that equals 20 (19 FAH categories plus a numéraire). This variable will address an endogeneity problem in unit values. More discussion of the censored data and price index is in Section 3.4.

The variable y_h represents the real total household expenditure with L being the highest degree of total expenditure polynomial to be determined by statistical tests, which is related to the second feature of the EASI model mentioned earlier. Following Lewbel and Pendakur (2009), we construct y_h as the Stone price-deflated total household expenditure:

$\log x_h - \sum_{j=1}^{J} w_{hj} \log p_{hj}$, where x_h is nominal total household expenditures on food and other goods and services. We also include interaction terms between the real total household expenditure and the price indexes to examine how Hicksian demand changes conditional on total expenditure.

Lastly, z_{hk} are K exogenous demand shifters that include household demographic variables. The household demographic variables control for observed taste differences among households, which includes log household head age, log household size, a martial status dummy (married vs. all other), household head education, maximum years of schooling within a household, and nine variables for the proportion of household members within each of the ten gender-specific age groups: 0-14, 15-29, 30-44, 45-64, and 65+, with the female 65+ age group set as the reference group.

Panel Model

Taking equation (3.1) as the basic model, we extend the model to a panel structure where time-invariant unobserved regional characteristics are controlled for. Following Meyerhoefer et al. (2005), we employ a correlated random effects specification by adding EA level means of the price index (p_{hj}) and EA level means of the interactions between the price index and the real total household expenditure ($y_h \times \log p_{hj}$). One may argue that household level means of these variables better capture household level heterogeneity. Adding the household level means, however, removes much of the price variations in the data, rendering it difficult to estimate the parameters. Similarly, we could not use differentiated EAs attached to each household who can be either original households surveyed at the first wave (termed as mother households) or their split-off households over waves. Instead, we track EAs for corresponding mother households of split-off households and assign those EAs to the split-off households if their EAs differ from their mother households' EA. The use of these newly defined EAs is justified under the assumption that split-off households continues to share similar characteristics (e.g., tastes) with their mother households (Atkin, 2013). These EA level mean variables are considered as additional demographic variables contained in demand shifter z_{hk}. Therefore, the number of demand shifters K increases by 40 in the panel model compared to the cross-sectional model.

3.4 Estimation Strategy

This section begins by reviewing the econometric techniques used to handle censored data as well as potentially endogenous variables in the model, and then describes a method to obtain nutrient elasticities.

3.4.1 Censored Demand

Following the vast literature on censored demand (Perali and Chavas, 2000; Meyerhoefer et al., 2005; Kasteridis et al., 2011), we use the Tobit model to characterize censoring. The Tobit model is less structural than the virtual price approach to censored demand. However, the latter approach is computationally infeasible for large systems because of multiple integrals of the large number of censored demand regimes. Also, the virtual price approach is reported to only work with the translog demand (Christensen et al., 1975), which has a less flexible functional form than EASI. Finally, it is unclear how the virtual price approach accounts for endogeneity in the explanatory variables, which may be a serious issue in micro data.

To estimate Tobit demand system (3.1), we use the extended Amemiya's generalized least squares (AGLS) estimator developed by Zhen et al. (2014) while controlling for Stone price-deflated total expenditure and price endogeneity (explained in more detail in the next subsection). The extended AGLS estimator builds on the standard AGLS estimator for single-equation limited dependent variable models and extends it to the context of a system of limited dependent variable equation. The estimator works in three steps. First, the reduced-form Tobit regressions are estimated equation-by-equation, where censored budget shares are the dependent variables. The explanatory variables are the exogenous demand shifters, instrumental variables, and residuals from least squares auxiliary regressions of endogenous prices on all exogenous variables and instruments. Second, the structural parameters of the budget share equation (3.1) are recovered using minimum distance (Wooldridge, 2010) and the correct asymptotic covariance matrix for the structural parameters. This procedure accounts for the correlation between the Tobit equations and the linear auxiliary regressions. Third, the minimum distance estimator is applied again to impose the utility-theoretic restrictions of homogeneity ($\sum_{j=1}^{J} \alpha_{ij} = 0$ and $\sum_{j=1}^{J} \alpha_{ijy} = 0 \ \forall i$)[5] and symmetry ($\alpha_{ij} = \alpha_{ji}$ and $\alpha_{ijy} = \alpha_{jiy}$)[6] on the latent demand. The three-step extended AGLS estimator is efficient among a class of limited information estimators (Newey, 1987). In comparison with full information maximum likelihood estimators that estimate all Tobit equations simultaneously (e.g., Dong et al. 2004), the extended AGLS is more feasible for estimating large demand systems, especially when some explanatory variables may be endogenous. Once we estimate demand system parameters, we can calculate predicted budget shares from which price and total expenditure elasticities are obtained. The derivation of the food demand elasticities is found in Appendix C.

3.4.2 Endogeneous Regressors

There are two sources of endogeneity in demand equation (3.1). First, $\log x_h$ is deflated by a Stone price index, which introduces budget shares into log real total expenditure y_h. We can easily correct this form of endogeneity by using category-level average of budget shares, ($\overline{w_j}$) to instrument w_{hj} in y_h.[7] The second form of endogeneity is concerned with the use of unit values as prices that are calculated by dividing expenditure by physical quantity. Although using micro-level data largely rules out demand-supply simultaneity common in aggregate data, product quality and price search behavior may cause unit values and prices paid to be endogenously determined with demand, respectively.

[5]Homogeneity restrictions make Marshallian demand homogenous of degree zero in price and expenditure.

[6]This restriction is for imposing Slutsky symmetry.

[7]Also, this form of endogeneity has been found to have little impact on empirical analysis in the Canadian and U.S. contexts (Lewbel and Pendakur, 2009; Zhen et al., 2014).

Unit values contain information on market prices and quality, which is demonstrated by the evidence of higher unit values for the top quartile households in Table 3.1. If one uses unit values without accounting for quality to estimate demand, the results will be biased (Cox and Wohlgenant, 1986; Deaton, 1988). In addition to controlling for quality, we should account for household cost minimization behavior in demand estimation. Households who actively search for lower prices are likely to exhibit different characteristics, which are unobserved to the econometrician, from those who do not search. The correlation between household cost minimization behavior and unobserved heterogeneity could cause biased estimates if left unaccounted for.

We address the unit value bias and potential biases from household price search behavior in two ways. First, we construct household Fisher Ideal price indexes at the food category level using food-item level unit values as elements. Specifically, the Fisher Ideal price indexes for household h, FAH category j $(j = 1, \ldots, J-1)$ is calculated as

$$p_{hj} = \sqrt{\frac{\sum p_{kh} q_{k0}}{\sum p_{k0} q_{k0}} \frac{\sum p_{kh} q_{kh}}{\sum p_{k0} q_{kh}}} \tag{3.2}$$

where p_{kh} and q_{kh} are the unit value and physical quantity (in kilogram, litre, and piece[8]) of food-item code k in FAH category j that household h reports, respectively. p_{k0} and q_{k0} are the base unit value and quantity of k calculated as averages of the 7 day-recall values of each, across sample households. The food-item level unit values are missing if the household did not consume the product during the 7 day-recall. We impute the missing unit values using median purchasing prices at various levels that are as much differentiated as the combination of food-item code, food unit, urban type, region, and EA and as simple as the combination of food-item code and food unit. We first replace missing unit values with the most differentiated-level median prices and sequentially proceed to the next less differentiated-level median prices until all missing values are filled.

The Fisher Ideal price index is superlative in that it is consistent with a second-order approximation to an arbitrary twice-continuously differentiable linear homogenous consumer

[8]The food group "eggs" is the only category that was reported in pieces.

cost function. Therefore, compared with using food-item level unit values as the price variables for equations (3.1), the Fisher Ideal index reduces the part of unit value bias due to within-category substitution. However, to the extent that the Fisher Ideal price index is constructed by food-item level unit values and does not address household unobserved characteristics related to price search behavior, it is still subject to endogeneity problems.

Therefore, we create instrument variables for Fisher Ideal price indexes. For Fisher Ideal price index p_{hj} for each household h and food category j, we calculate a mean price index using price index values of other households (donor households) in the same region and the survey date (month and year). If a household has no such donor household, we use a mean price index calculated over other households in the same region and the survey wave for the household.

The price index for the numéraire good is the Tanzanian consumer price index (CPI) less food, alcoholic beverages, tobacco, and narcotics. The instrument for the numéraire good is the CPI lagged by two months.[9] Identification of the price coefficients relies on the spatial and temporal variations in the price indexes for 19 FAH category and the numéraire (Deaton, 1988; Beatty, 2010).

3.4.3 Nutrient Elasticities Derivation

We follow Huang (1996) and Huang and Lin (2000) to estimate nutrient elasticities in a food demand system. The fundamental premise is that nutrient intake largely depends on food consumption and that food consumption is affected by individuals' sensitivity to changes in prices and income. According to this notion, we retrieve the calorie and nutrient elasticities using price and total expenditure elasticities. The derivation of nutrient elasticities is as follows.

[9] As Tanzanian CPI is not available from September 2008 to Septermber 2009, we use the CPI of the month from which data is available for the missing instrument variable for the price index for the numéraire good and two months ahead of that month for the price index for the numéraire good.

We start with the following total quantity of nutrient γ (φ_γ) from consumption of each food group i:

$$\varphi_\gamma = \sum_i a_{\gamma i} Q_i(p_1, ..., p_n, m). \tag{3.3}$$

where $a_{\gamma i}$ is the quantity of nutrient γ per unit of food category i and Q_i is the total consumption quantity of food category i which is a function of prices of the food and other goods along with total expenditure m.

We then totally differentiate equation (3.3) with respect to prices and total expenditure, which results in:

$$\frac{d\varphi_\gamma}{\varphi_\gamma} = \sum_i a_{\gamma i}\, dQ_i \frac{1}{\varphi_\gamma} \tag{3.4}$$

As mentioned, we use price and total expenditure elasticities to back out nutrient elasticities. The term dQ_i in equation (3.4) has the link. When the food demand equation Q_i is totally differentiated with respect to prices and total expenditure, we have the following formula.

$$\begin{aligned} \frac{dQ_i}{Q_i} &= \sum_j e_{ij} \frac{dp_j}{p_j} + \eta_i \frac{dm}{m} \\ dQ_i &= [\sum_j e_{ij} \frac{dp_j}{p_j} + \eta_i \frac{dm}{m}] Q_i \end{aligned} \tag{3.5}$$

where e_{ij} indicates own- or cross-price elasticities and η_i represents total expenditure elasticities. We subsitute the formula of dQ_i in (3.5) into equation (3.4), which yields:

$$\begin{aligned} \frac{d\varphi_\gamma}{\varphi_\gamma} &= \sum_i a_{\gamma i} [\sum_j e_{ij} \frac{dp_j}{p_j} + \eta_i \frac{dm}{m}] \frac{Q_i}{\varphi_\gamma} \\ &= \sum_j \pi_{\gamma j} \frac{dp_j}{p_j} + \rho_\gamma \frac{dm}{m} \end{aligned} \tag{3.6}$$

where $\pi_{\gamma j} = \sum_i a_{\gamma i} Q_i \frac{e_{ij}}{\varphi_\gamma}$ is the elasticity of demand for nutrient γ with respect to price of the j th food and $\rho_\gamma = \sum_i a_{\gamma i} Q_i \frac{\eta_i}{\varphi_\gamma}$ is the total expenditure elasticity of demand for nutrient γ.

3.5 Results

3.5.1 Test Statistics

As explained in Section 3.4, we estimate the system of $J-1$ Tobit equations of (3.1) using the extended AGLS. The parameters of the budget share equation for the numéraire good, which is not censored, are recovered at post-estimation using the homogeneity, symmetry, and adding-up ($\sum_{i=1}^{J}\mu_j = 1$ and $\sum_{j=1}^{J}\beta_{jr} = 0$) restrictions[10] on the latent demand. We first examine a shape of Engel-curve that best fits into the data by testing the joint significance of coefficients β_{iL} ($i = 1, \ldots, J-1$) by minimum distance as increasing the value of L from 1. We find that the most supported degree of polynomial on real total household expenditure is $L = 3$ both for cross-sectional and panel models.[11] When conducting this test, we do not impose the homogeneity and symmetry conditions on the demand system. Otherwise, the test would become a joint test of β_{iL} and these economic restrictions.

We also test for the joint significance of coefficients α_{ijy} on the interaction between log price and real income. Without imposing the symmetry and homogeneity conditions, the test produces a test statistic of 2906.54 (p-value $<$ 0.000; cross-sectional model) and 2674.66 (p-value $<$ 0.000; panel model) with 380 degrees of freedom. This result reinforces the superiority of the two-way EASI model over other conventional demand systems by allowing the Hicksian demand to vary with total expenditures.

[10] Adding-up restrictions ensure that the estimated budget shares sum to one. Satisfying with the above three parametric restrictions, a set of demand functions are *integrable*.

[11] Under the null that ${y_h}^L$ can be excluded from the demand system, the test statistic is asymptotically distributed as $\chi^2(J-1)$. When $L = 3$, the test statistics for cross-sectional and panel models are 221.90 and 185.40, respectively, while they are 62.88 and 64.05 when $L = 4$. Although the null with the test statistics for $L = 4$ is not rejected, we find that adding additional polynomial term only creates numerical issues due to increasing collinearity. Under the model with $L = 4$, the signs and magnitudes of the expenditure elasticities on some food groups (including staple foods) are unreasonably large and some are negative even for the poorest households. Also, gaps between elasticities at mean and median are substantially large. Therefore, we decide to include upto the cubic real expenditure term in the model.

3.5.2 Price Elasticities

As discussed in Section 3.3, we estimate two EASI models: one is a basic cross-sectional model and the other is a panel model. Given that the latter model controls for regional unobserved heterogeneity at the EA level but the former model does not, we start with comparison of price elasticities results between two models. We report the Marshallian price elasticities at median for the whole sample in Table 3.2 for the cross-sectional model and in Table 3.3 for the panel model.

We highlight two noticeable differences between the results from the two models. First, the absolute magnitudes of own-price elasticities estimated from the panel model tend to be smaller than those from the cross-sectional model. Figure 3.2 shows this finding more clearly. This result provides evidence that neglecting the effects of regional food tastes on price variations could bias own-price elasticities upward in absolute terms. For example, suppose region A where households have traditionally allocated a large portion of their expenditure on rice has lower rice price than region B where households do not exhibit strong preferences for rice.[12] When the price of rice increases by the same rate, the rice demand in region A responds negatively to this shock, but to a lesser extent compared to region B because of the strong local tastes for rice and the smaller absolute rise in price. In other words, the magnitude of the change in rice demand is contingent upon not only pure price effects but also differences in regional preferences. Therefore, if we fail to address the contribution of the latter part in the model, we are likely to overestimate price sensitivity of the quantity consumed of a food at the aggregated level. The second difference is that the cross-sectional model reports more cross-price elasticity pairs that are statistically significant when compared to the panel model, implying that ignoring unobserved regional heterogeneity is likely to lead over-rejection of the null effect of relative variations in food prices on consumption.

[12] Over the generations, households have evolved their preferences for local foods that are readily available in their region and easy to grow. The easy access to the local foods results in low price of these foods, enhancing the households' strong tastes for the foods again (See Atkin 2013).

Both cross sectional and panel results consistently show that households' food consumption is more sensitive to own-price variations compared to cross-price changes. For the whole sample, the demand for rice, wheat and other cereals, roots, tubers, and other starches, nuts and seeds, and poultry is highly responsive to own-price changes with the values of the elasticities close to or over two. By contrast, the demand for sugars, vegetables, eggs, fats and oils is relatively less responsive to own-price variations. Among the cross-price elasticity pairs for major food groups that are statistically significant in both results, substitutive relationship includes rice and pulses; rice and dairy; maize and sugar; cassava and red meat; cassava and dairy; roots, tubers, and other starches and wheat and other cereals; roots, tubers, and other starches and vegetables; roots, tubers, and other starches and dairy; sugar and pulses; pulses and vegetables; pulses and eggs; pulses and fish and seafood; vegetables and fish and seafood; and red meat and eggs. Complementary relationship includes wheat and other cereals and red meat; cassava and roots, tubers, and other starches; roots, tubers, and other starches and pulses; vegetables and eggs; vegetables and dairy; red meat and dairy; and poultry and eggs.

We observe a few cases where the signs of t-values are inconsistent with those of the corresponding estimated elasticities in both tables. Because we calculate t-values at median using simulated elasticities from 100 random draws from a multivariate normal distribution of the model parameters with the mean vector and and their variance-covariance matrices, this inconsistency could happen when a distribution of the simulated elasticities is not well overlapped with the distribution of the elasticities from the actual data. We do not believe this is problematic as the discrepancies between the estimated elasticities and the corresponding t-values (mostly small in magnitude) are marginal. Overall, both cross sectional and panel results provide a similar qualitative explanation of households' food consumption patterns in response to food price changes but the panel model certainly yields more credible estimates of the elasticities by accounting for unobserved regional differences.

Further, it is worth exploring how households in lower per capita quartiles behave differently from those in higher per-capita quartiles when facing food price changes. Focusing on panel model results, we report own-price elasticities by quartile in Table 3.4 and relegate the whole quartile-specific price elasticity results to Appendix C. We first find that the magnitudes of own-price elasticities generally fall in absolute sense with rising per-capita total expenditure, suggesting that households become more price inelastic as their income is higher. This trend is prominent among rice, wheat and other cereals, and poultry. For instance, 1% increase in poultry price reduces poultry consumption by 3.10% in the bottom quartile but the same rate of the price rise only decreases the demand by 0.46% in the top quartile. We also observe that own-price elasticities increase in rising per-capita total expenditure among cassava, vegetables, dairy, coffee, tea, and cocoa, and other food. Possible reasons to explain this finding include higher income households' lower preferences on these food groups or abundant consumption of them.

For the poorest households, almost all food groups except sugar, vegetables, eggs, dairy, fats and oils, coffee, tea, and cocoa, and other food are own-price elastic. In particular, the own-price elasticities for rice, wheat and other cereals, roots, tubers, and other starches, nuts and seeds, and poultry are substantially large. Among all quartiles, the own-price elasticity for maize (the primary staple food in Tanzania) is statistically significant with the exception of the top quartile households. This indicates that stabilization of maize price benefits most Tanzanians except the top 25% rich.

Comparing cross-price elasticities across quartiles, we find that poorer households demonstrate more food substitutes. Specifically, there are 28 significant substitutes between food groups among the poorest households whereas there are only 6 substitutes found for the richest households. This finding is mainly due to higher vulnerability to food price variations among the poor but partly due to dynamics of food preferences across quartiles. For instance, for the bottom quartile households, rice, roots, tubers, and other starches, and pulses are good substitutes to vegetables, whereas no such substitutes exist for vegetables for the top quartile households causing the highest own-price elasticity in this quartile. Also, poor households appear to favor pulses as pulses are a good substitute for not only vegetables but also for animal products and fish and seafood. On the contrary, pulses are not substituted with any other food categories among the richest households.

3.5.3 Total Expenditure Elasticities

Here, we first compare results of total expenditure elasticities at median between two models and then closely examine the panel result to further elucidate Tanzanians' food demand response to total expenditure variation by per-capita expenditure quartile. The last column in Table 3.2 shows the cross-sectional result and the last column in Table 3.3 presents the panel result for the whole sample.

Both models suggest that expenditure elasticities are generally greater than price elasticities. This result is consistent with previous research showing that households' food consumption patterns are more sensitive to income effects than price effects (Weliwita et al., 2002; Abdulai and Aubert, 2004). Unlike the price expenditure elasticity results, we do not find that cross-sectional results provide larger magnitudes of the total expenditure elasticities. This result implies that unobserved regional heterogeneity is more important in price sensitivity than total expenditure sensitivity of the food consumption.[13]

Despite slight differences in magnitudes, main findings of the results from the two models are analogous. Total expenditure elasticities for cheap staple foods, such as maize and cassava, are relatively small and insignificant. This suggests that households tend to shift away

[13]Total expenditure elasticities estimation still benefits from addressing regional heterogeneity in that it gains more precision. Continuing the rice example in region A and B, we can explain how households in two regions behave differently in terms of rice demand when they have additional income. In region A with rice-loving households, the quantity consumed would increase with rising total expenditure but decrease in relative terms if the households already consume enough rice. On the other hand, households in region B would exhibit an overproportional consumption increase with rising total expenditure if rice is a luxury good there. This example shows that addressing the heterogeneity certainly helps better understand income effects on food consumption.

from the satiated food groups with additional income. By contrast, total expenditure elasticities for red meat, poultry, and eggs are over two (panel) or close to two (cross-sectional), implying that the quantities consumed of these animal source products almost double with 1% increases in total expenditures. This finding contrasts with the finding of Weliwita et al. (2002) showing that demand for meat is inelastic with respect to the expenditure on *food* through a conditional food demand estimation. Among non-staple foods excluding "other food", an increase in total expenditure leads to a less than a proportionate increase in demand for vegetables, fruit, fish and seafood in both models with fats and oils being added in the panel model only. Because the average food budget share on these categories varies by quartile (see Table 3.1), we expect that perceiving them as necessary goods would not hold across all per-capita total expenditure groups. We obtain better insight on total expenditure effects on food consumption by examining total expenditure elasticities by per-capita total expenditure quartile.

Figure 3.3 shows quartile-specific elasticities of quantity consumed with respect to total expenditure changes according to the panel results. We observe notable differences in total expenditure effects on food consumption across quartiles. The food consumption sensitivity to a marginal increase in total expenditure tends to decline with increases in households' per-capita total expenditures. For instance, the number of food categories with the expenditure elasticities greater than two(one) is 8(12) in the bottom quartile as compared to 0(9) in the top quartile.

As expected, fish and seafood are luxury goods (with total expenditure elasticity of 1.37) for the bottom quartile households; a result that does not hold for the other groups. Expenditure elasticity for fruit is 2.21 among the poorest households but it drops to 0.94 in the second quartile. Moreover, the lowest quartile households substantially increase their consumption on poultry and eggs relative to other quartile groups. With 1% increase in total expenditure, their consumption on these food categories increases by 4.32% and 3.75%, respectively. These results imply that the poorest households have strong preferences for these products, yet their desired consumption is not successfully satisfied given restrictions associated with their current income. Lastly, although statistically insignificant, total expenditure elasticities for cassava are negative for the third and fourth quartiles (-0.01 and -0.09, respectively). Combined with the relatively large own-price elasticities for cassava among these households, this result provides suggestive evidence that relatively richer households could perceive cassava as inferior food.

3.5.4 Nutrient Elasticities

For nutrient elasticities, we only focus on panel results given that these values are calculated from price and expenditure elasticities that are estimated more accurately using the panel model. Table 3.5 presents nutrient elasticities with respect to food prices and total expenditure for the whole sample. For most food categories, nutrient price elasticities are small and insignificant, suggesting that nutrient consumption is price-inelastic. By contrast, nutrient total expenditure elasticities are fairly large and all significant at the 1% level. These findings correspond with the larger total expenditure effects on food consumption, placing more weight on income support policies for nutrition improvement (Ecker and Qaim, 2011).

The nutrient price elasticities result for the aggregated sample indicates that the price of maize and pulses is considered paramount for Tanzanians' overall nutrients intake. An increase in the price of these foods negatively affects most nutrients with the exception to fat, sugar, and vitamin A. As for vitamin A, the price of roots, tubers, and other starches is the most important determinant. When the price of the food category increases by 1%, consumption of vitamin A declines by 0.84%. We also find that substitution effects among food groups play an important role determining relative changes in nutrients. For instance, an increase in the price of maize is associated with higher carbohydrate (sugar) intake, implying that households substitute maize with carbohydrate-rich food, such as sugar. Indeed, maize and sugar are substitutes shown in both Table 3.2 and Table 3.3. In another example, we find positive effects of an increase in the price of coffee, tea, and cocoa on most macro- and micro-nutrients consumption. This phenomenon is also driven by its substitution effects with many food groups that include maize, wheat and cereals, pulses, vegetables, red meat, and fats and oils.

We also calculate nutrient price elasticities by per-capita total expenditure quartile. For the sake of clarity, we plot the elasticities by nutrient and quartile in Figure 3.4 to Figure 3.7 with significance denoted at the 5% level. We first find that nutrient impacts of price variations are greater and more extensive for the poorer households, suggesting that food price-related policies would be particularly effective when they target the poor. This is perhaps an expected result given that poor households are likely to be more vulnerable to increases in food prices. For instance, the price of maize, pulses, nuts and seeds have the strongest effects on both calorie intake and the consumption of macro- and micro-nutrients for the bottom 25% of households. The effects of these food groups, however, diminish with rising total expenditure.

Consistent with the result from the whole sample, we also find that the price of roots, tubers, and other starches is important for vitamin A intake, especially for the poorest households. A 1% increase in its price results in a drop in vitamin A intake by 1.26% with the magnitude of the elasticity decreasing in absolute terms with rising income. For the top 50% of households, the price of vegetables becomes important for consumption of vitamin A. In addition, our finding of no significant nutritional impacts of the price of poultry is interesting because households, especially the poor, exhibit significantly own-price elastic demand for this food category. This can be the case when households substitute poultry with other cheaper and nutrient-dense foods so that any decrease in poultry consumption does not necessarily reduce households' overall diet quality.

We further explore how nutrient total expenditure elasticities differ by per-capita total expenditure quartile, which is shown in Figure 3.8. Nutrient intake elasticities with respect to total expenditure generally increase as per-capita total expenditure rises for most nutrients, with the exception of fat and sugar. With additional income, the poorest households increase fat intake the most while increase vitamin A intake the least among four quartiles. Indeed, the marginal change in micronutrient intake for the bottom quartile households is the smallest among all quartiles when additional income is given. For instance, the poorest households only increase intake of vitamin A and total folate by 0.65% and 0.78%, respectively, with a 1% increase in total expenditure, whereas the magnitudes of the same micronutrients intake are 1.14% and 0.89% for the third quartile and 0.99% and 0.89% for the fourth quartile households. Combined with the evidence that micronutrient deficiencies are particularly significant in Southeast Asia and Sub-Saharan Africa (Ramakrishnan, 2002), our results suggest that greater efforts and attention are needed to improve micronutrient malnutrition of the poor in Tanzania.

3.6 Discussion and Conclusions

In this article, we provide new estimates of households' food and nutrient consumption sensitivity to food price and total expenditure changes by controlling for regional price variations attributed by unobserved regional characteristics. Using the Tanzanian LSMS data, we first estimate a cross-sectional two-way EASI demand system with 19 food categories and a numéraire good that allows flexible Engel curves and price effects in the data while addressing censoring of the reported consumption. We then successfully extend the model to a panel structure with a correlated random effects specification to account for the regional heterogeneity at the EA level. Using food consumption elasticities with respect to food prices and total expenditure which are estimated from the model parameters, we derive nutrient price and total expenditure elasticities for calorie and 7 key nutrients at the aggregated level and four per-capita total expenditures levels.

Comparing estimated elasticities between the two models, we first find that failing to account for the unobserved heterogeneity is likely to lead upwardly biased own-price elasticities in absolute sense and over-rejection of the null cross-price effects. We also provide evidence that effects of total expenditure on food consumption, mostly significant, are greater than price effects. We find that effects of food price and total expenditure vary by per-capita total expenditure. Both for changes in food and nutrient consumption, poorer households become more sensitive to increases in food prices. Whereas for changes in food consumption, poorer households are more responsive to income growth, for changes in nutrient consumption they are less responsive to income growth. Another notable finding is that from a nutritional balance perspectives, households in poverty poorly diversify their diets with additional income, resulting in micronutrient malnutrition.

Our results have a number of broader policy implications for nutrition improvement. First, income support policies may be generally more effective than price policies in changing the level of consumption for overall food groups, and therefore overall diet quality. Yet, given that the poorest are the biggest victims of food price fluctuations, price regulations or price subsidy programs for staple foods, if designed with great care, would help nutrition improvement for the poor. Second, nutrition education programs would play a positive role in inducing micronutrient-dense food choices of the poor households. Additional nutritional interventions such as direct food transfer programs could benefit the poor by helping form the habit of consuming micronutrient-rich foods as well as readily access those foods.

Lastly, because our findings are based on the 7-day food consumption recalls from household surveys, measurement errors in the quantities consumed and therefore nutrient consumption amounts are unavoidable. Also, the data restrict us from including food-away-from-home (FAFH) in the analysis although FAFH is becoming more and more important especially for wealthier households in urban area (Ma et al., 2006; Gäl et al., 2007). We are limited to analyzing food and nutrient intake at the household level as opposed to at the individual household member level, assuming that no competition exists among members of the household for available foods. These limitations which are not uncommon in other similar studies reveal a future research direction that is conditional on data availability. It would be informative to incorporate detailed FAFH consumption into food demand system analysis and examine nutrition transition driven by FAFH among urban households in developing countries.

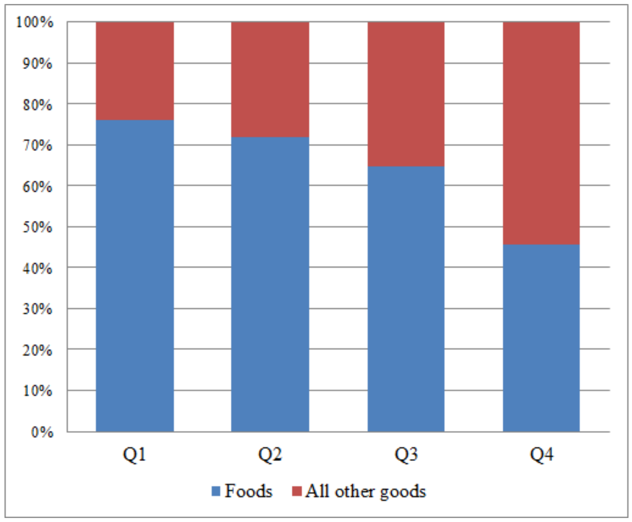


Figure 3.1: The Budget Share on Food by Per-capita Total Expenditure Quartile

Notes: This figure shows how households' budget share on foods varies by per-capita total expenditure quartile.

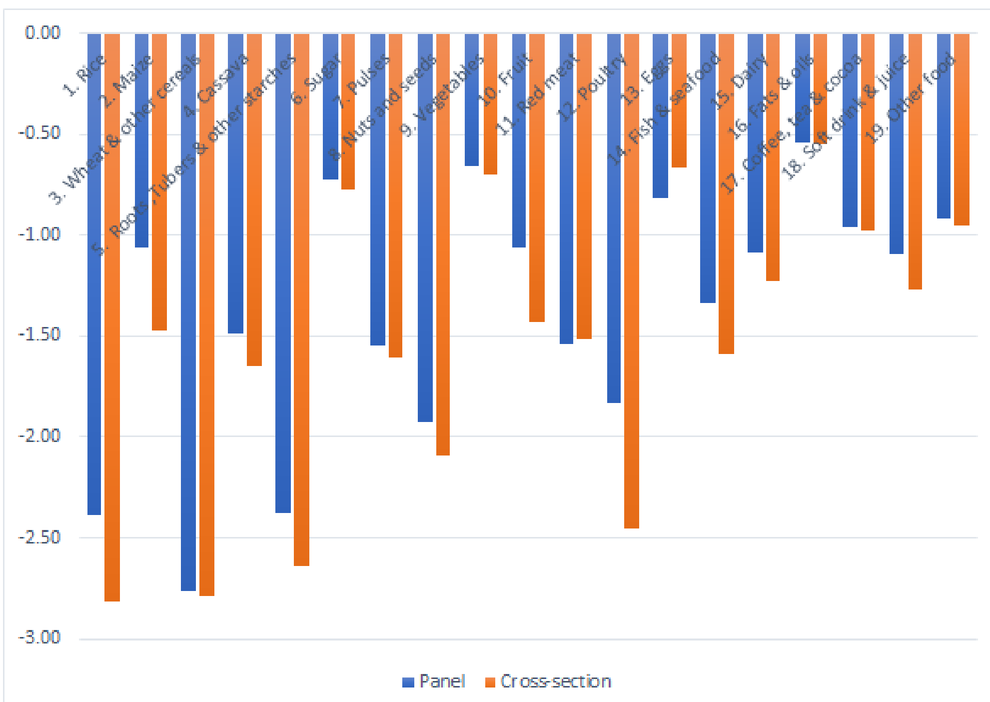


Figure 3.2: Own-Price Elasticities: Panel vs. Cross-Section

Notes: This figure shows own-price elasticities, one estimated from the panel and the other estimated from the cross-sectional model.

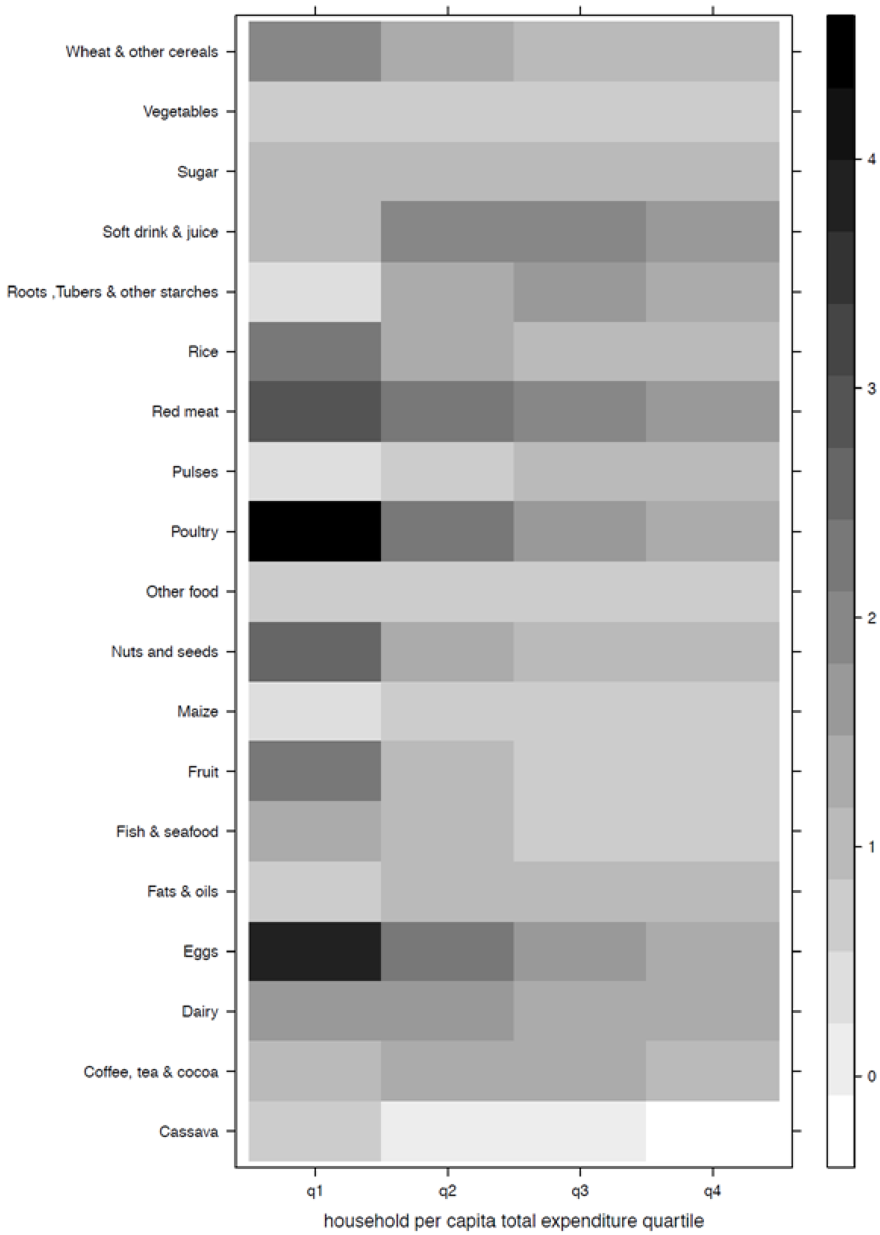


Figure 3.3: Total Expenditure Elasticities of Food Demand by Quartile

Notes: This figure shows quartile-specific percent changes of food consumption (kg/litre/piece) of households when household total expenditure increases by 1%. The darker the box is, the more quantity of the food households consume with additional income.

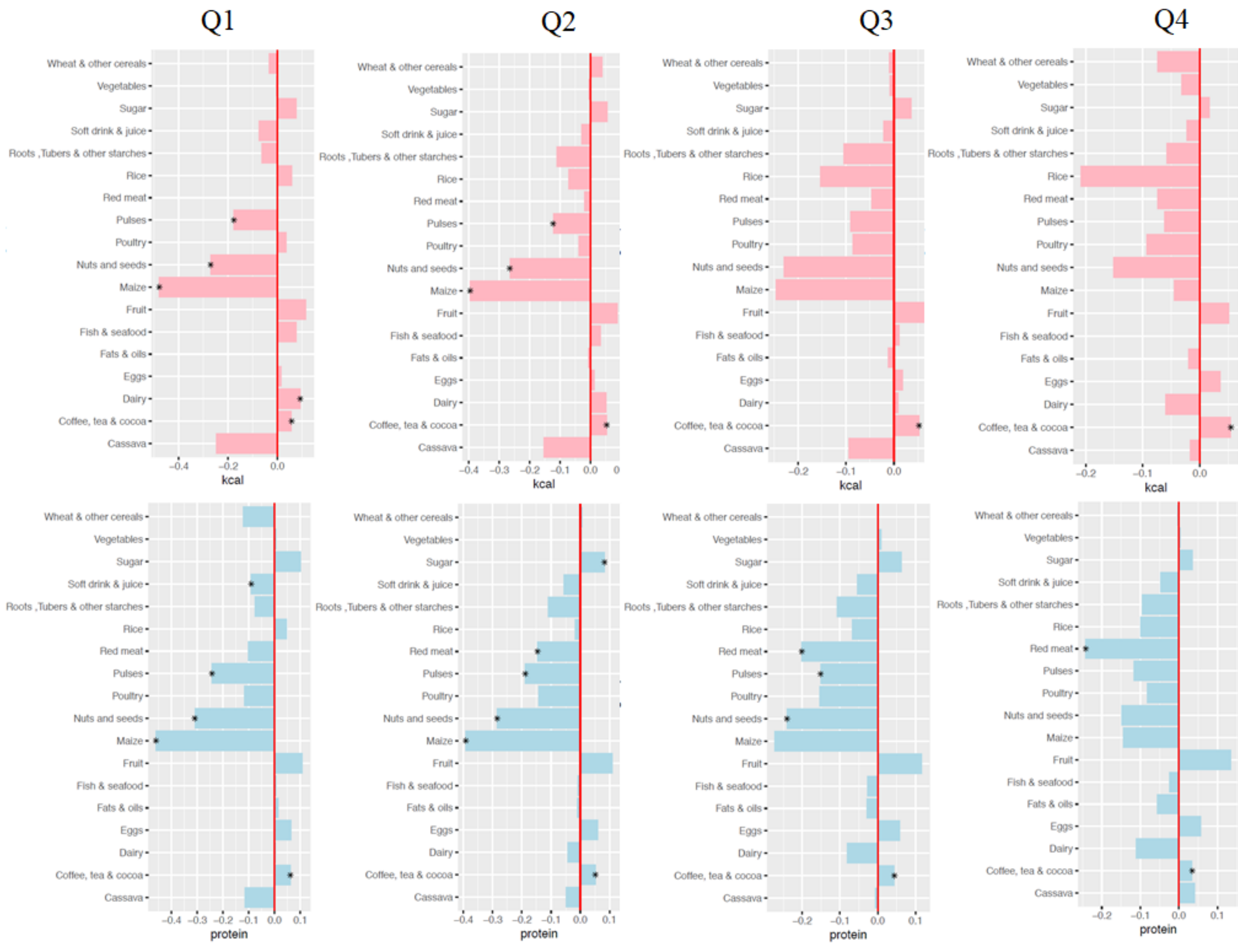


Figure 3.4: Nutrient Price Elasticities by Quartile: Energy and Protein

Notes: This figure shows quartile-specific percent changes of nutrient (kcal and protein) intake of households when food price increases by 1%.

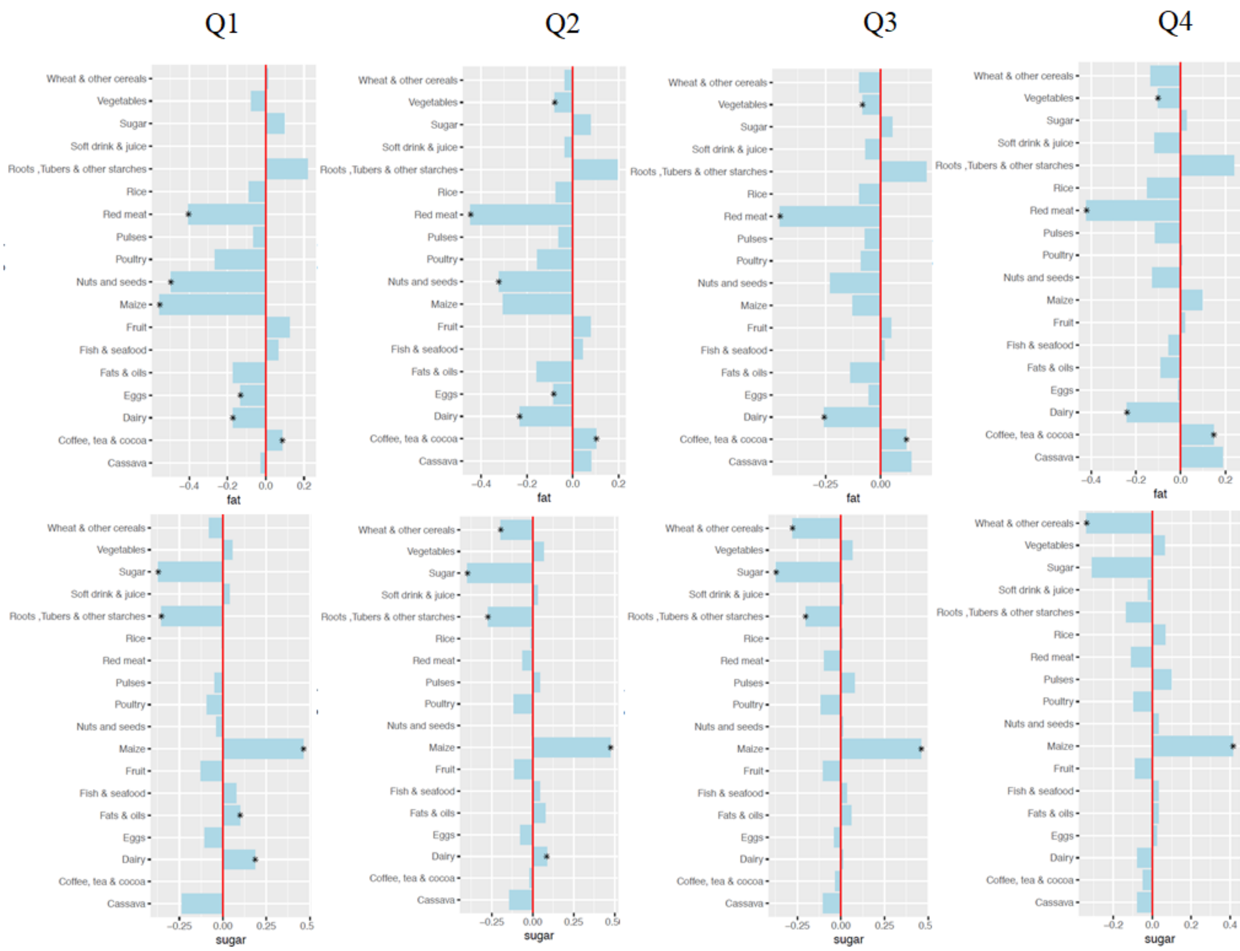


Figure 3.5: Nutrient Price Elasticities by Quartile: Fat and Sugar

Notes: This figure shows quartile-specific percent changes of nutrient (fat and sugar) intake of households when food price increases by 1%.

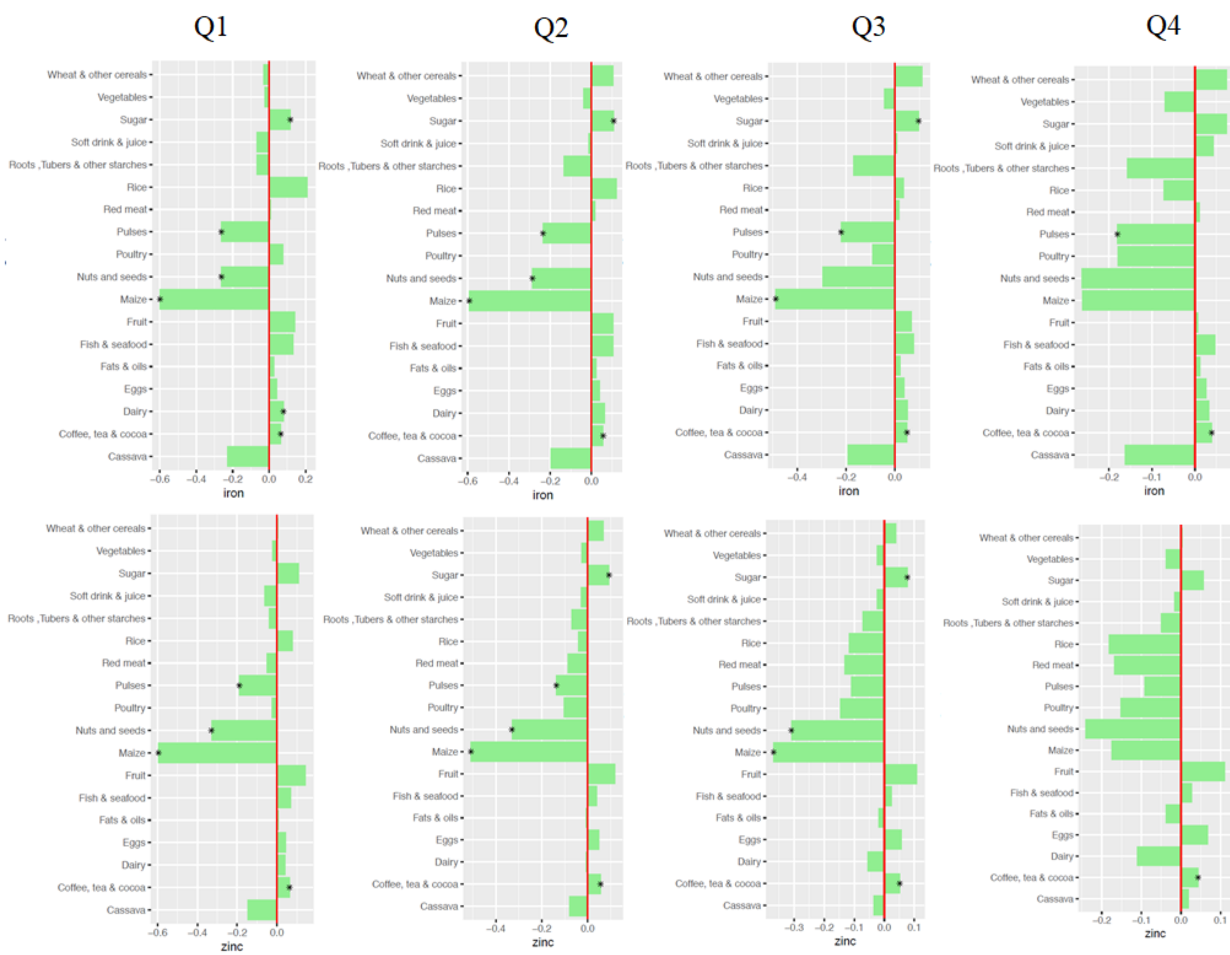


Figure 3.6: Nutrient Price Elasticities by Quartile: Iron and Zinc

Notes: This figure shows quartile-specific percent changes of nutrient (iron and zinc) intake of households when food price increases by 1%.

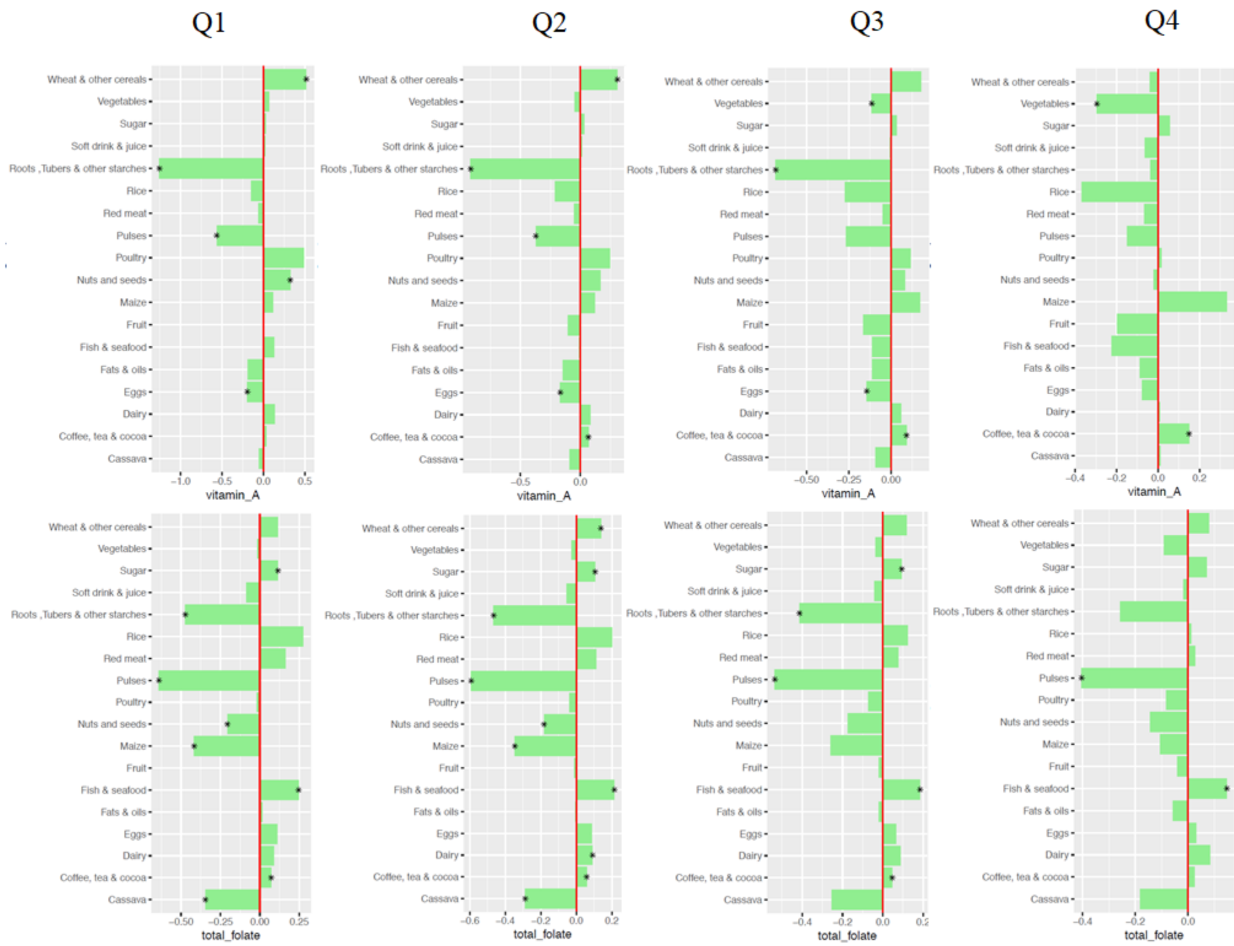


Figure 3.7: Nutrient Price Elasticities by Quartile: Vitamin A and Total Folate

Notes: This figure shows quartile-specific percent changes of nutrient (vitamin A and total folate) intake of households when food price increases by 1%.

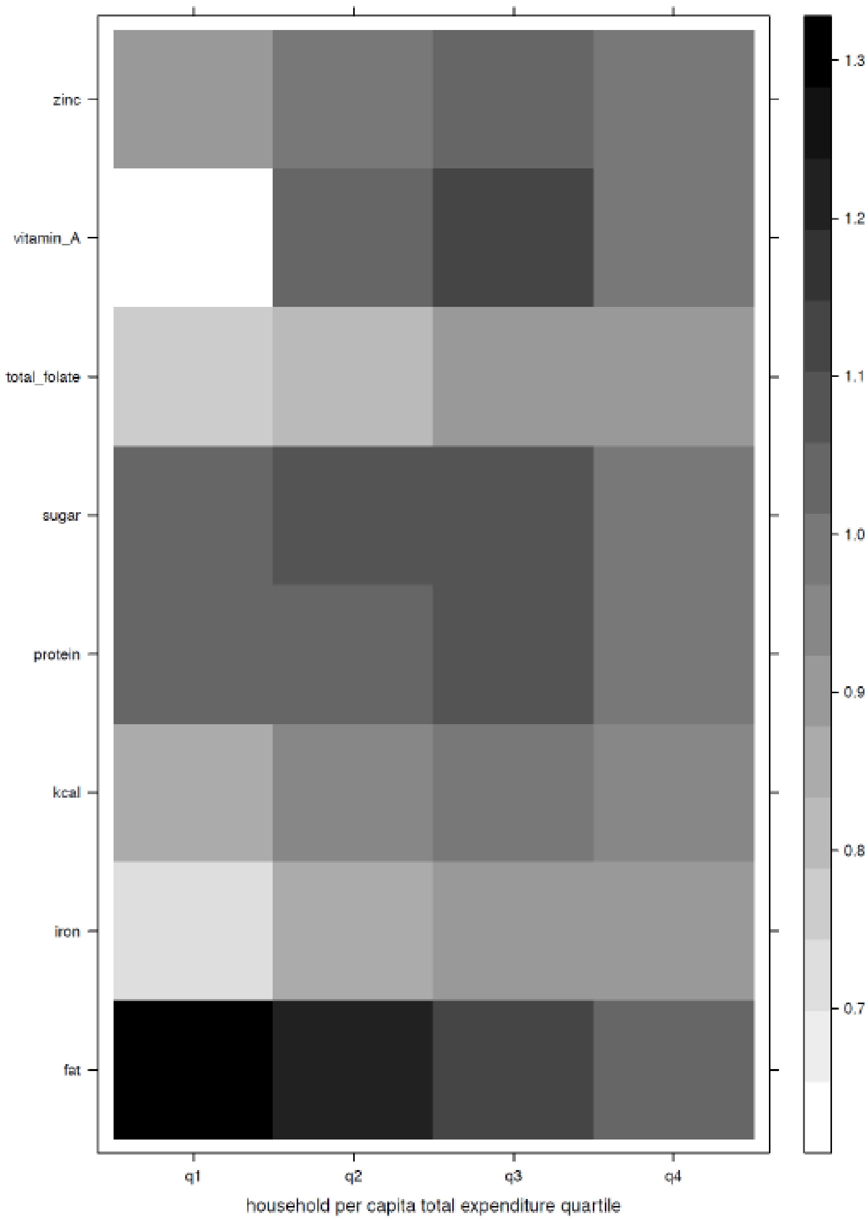


Figure 3.8: Total Expenditure Elasticities of Nutrient Intake by Quartile

Notes: This figure shows quartile-specific percent changes of nutrient intake of households when household total expenditure increases by 1%. The darker the box is, the more quantity of the nutrient households consume with additional income.

Table 3.1: Descriptive Statistics

	Average Food Budget Share				Average Unit Value			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Rice	0.072	0.113	0.127	0.125	1157	1222	1328	1513
Maize	0.259	0.211	0.166	0.103	606	743	859	960
Wheat and other cereals	0.042	0.047	0.062	0.077	1320	1536	1635	1690
Cassava	0.093	0.053	0.033	0.015	396	477	553	657
Roots, tubers, and other starches	0.062	0.072	0.073	0.061	472	598	710	862
Sugar	0.033	0.042	0.046	0.048	1711	1805	1854	1910
Pulses	0.079	0.061	0.053	0.039	1169	1297	1400	1562
Nuts and seeds	0.025	0.019	0.014	0.011	1321	1636	1861	2072
Vegetables	0.112	0.093	0.091	0.094	859	971	1061	1134
Fruit	0.039	0.053	0.065	0.072	635	767	863	1010
Red meat	0.029	0.051	0.071	0.107	3016	3639	4271	4792
Poultry	0.018	0.031	0.029	0.042	3197	3828	4165	5167
Eggs	0.003	0.004	0.006	0.011	201	214	233	259
Fish and seafood	0.055	0.063	0.071	0.074	2508	2854	3077	3787
Dairy	0.020	0.025	0.026	0.035	695	766	948	1813
Fats and oils	0.033	0.038	0.039	0.044	2577	2851	2864	2963
Coffee,tea, and cocoa	0.005	0.007	0.008	0.009	8951	9707	10197	10556
Soft drink and juice	0.007	0.008	0.012	0.026	547	765	1023	1360
Other food	0.013	0.008	0.008	0.007	787	792	889	1041

Notes: This table shows the summary statistics of the average budget share of the total food expenditure and average unit prices by food category and per-capita total expenditure quartile.

Table 3.2: Price and Total Expenditure Elasticities for the Whole Sample: Cross-Sectional Model

	Food Category																			Expenditure
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. Rice	-2.82*	1.00*	-0.21*	0.26	-0.11	-0.05	0.53*	-0.33*	0.20*	-0.25*	0.52*	-0.54*	-0.08	-0.61*	0.36*	-0.06	0.03	0.03	-0.04 *	1.18*
	(-6.20)	(3.25)	(-2.16)	(1.77)	(-1.03)	(-1.32)	(3.07)	(-2.46)	(2.70)	(-2.09)	(2.91)	(-2.50)	(-1.62)	(-3.29)	(2.81)	(-1.16)	(1.37)	(0.49)	(-2.51)	(4.86)
2. Maize	0.55*	-1.47*	0.47*	-0.11	-0.04	0.17*	-0.14	-0.32	-0.08	0.43*	-0.08	0.31	0.04	0.17	-0.05	0.02	0.06*	0.02	0.02*	0.78
	(2.37)	(-4.49)	(2.71)	(-0.90)	(-0.34)	(2.95)	(-1.75)	(-2.14)	(-1.94)	(2.10)	(-1.19)	(1.52)	(1.47)	(1.81)	(-1.17)	(0.74)	(2.79)	(0.65)	(2.45)	(1.83)
3. Wheat†	-0.30*	1.20*	-2.79*	0.22	1.08*	-0.03	0.12	0.49*	0.00	-0.43*	-0.42*	-0.12	0.03	-0.02	-0.06	0.13*	0.07*	-0.40*	-0.07*	1.30*
	(-2.54)	(4.80)	(-9.21)	(1.75)	(4.78)	(-0.67)	(1.56)	(3.58)	(0.13)	(-4.02)	(-3.13)	(-0.77)	(0.54)	(-0.52)	(-1.09)	(2.40)	(2.62)	(-4.08)	(-3.08)	(6.55)
4. Cassava	0.37*	-0.04	0.22*	-1.65*	-0.12	0.02	-0.23*	-0.25	0.07	-0.08	0.65*	-0.30	0.05	0.12	0.42*	0.14	0.00	-0.19	0.02	-0.07
	(2.37)	(-0.38)	(2.32)	(-5.29)	(-0.65)	(0.54)	(-2.07)	(-1.02)	(1.25)	(0.07)	(2.49)	(-0.97)	(1.12)	(1.31)	(2.53)	(1.87)	(-0.30)	(-1.94)	(1.90)	(1.19)
5. Roots†	-0.13	-0.07	0.79*	-0.19	-2.64*	-0.11*	-0.83*	0.55*	0.29*	0.07	0.16	0.63*	-0.08	-0.06	0.27*	0.04	-0.09*	-0.05	-0.01	1.15*
	(-1.16)	(-0.49)	(3.50)	(-1.27)	(-5.55)	(-2.20)	(-3.40)	(2.46)	(3.17)	(0.39)	(1.19)	(2.08)	(-1.54)	(-1.08)	(2.32)	(1.17)	(-2.94)	(-0.60)	(-0.91)	(3.33)
6. Sugar	-0.14	0.77*	-0.06	-0.03	-0.25*	-0.77*	0.38*	0.13	0.18	-0.18*	-0.07	0.02	-0.09	-0.03	0.09	0.21*	-0.07*	0.20*	0.01	1.01*
	(-1.43)	(5.87)	(-0.67)	(-0.42)	(-2.55)	(-5.99)	(2.93)	(1.58)	(2.71)	(-2.41)	(-0.89)	(0.06)	(-1.24)	(-0.62)	(1.79)	(3.14)	(-2.07)	(3.51)	(0.65)	(7.18)
7. Pulses	0.75*	-0.34*	0.14	-0.36*	-1.07*	0.20*	-1.60*	0.04	0.19*	0.04	0.24*	0.18	0.34*	0.68*	0.09	-0.15*	0.08*	-0.19*	-0.02	0.77*
	(3.22)	(-2.25)	(1.71)	(-2.35)	(-3.77)	(2.60)	(-5.95)	(0.25)	(2.57)	(0.55)	(2.25)	(1.21)	(3.01)	(3.86)	(1.20)	(-2.11)	(2.46)	(-2.37)	(-1.52)	(3.10)
8. Nuts†	-0.64*	-1.28*	0.66*	-0.48	1.05*	0.09	0.03	-2.09*	-0.06	0.22	-0.50	-1.78*	-0.36*	0.21	-0.41*	0.13	0.03	0.06	0.09*	1.21*
	(-2.01)	(-2.45)	(2.51)	(-1.63)	(2.21)	(1.38)	(-0.16)	(-4.97)	(-0.47)	(1.21)	(-1.91)	(-2.38)	(-2.30)	(1.52)	(-2.20)	(1.45)	(1.01)	(0.26)	(2.83)	(3.00)
9. Vegetables	0.30*	-0.18*	0.04	-0.01	0.38*	0.10*	0.18*	-0.02	-0.70*	-0.22*	-0.02	-0.08	-0.11*	0.25*	-0.08*	-0.05	0.08*	0.14*	-0.04*	0.72*
	(5.46)	(-3.08)	(1.17)	(0.38)	(5.38)	(3.11)	(3.13)	(-0.01)	(-13.18)	(-3.37)	(-0.51)	(-0.71)	(-2.87)	(6.18)	(-2.43)	(-1.69)	(6.23)	(3.75)	(-4.00)	(7.65)
10. Fruit	-0.39	1.44*	-0.49*	-0.17	0.15	-0.10	0.06	0.19	-0.28*	-1.43*	0.54*	-0.65	-0.17	-0.12	0.33*	0.04	0.00	0.18	-0.06*	0.75*
	(-1.40)	(2.18)	(-2.09)	(-0.58)	(0.68)	(-1.61)	(0.43)	(1.28)	(-2.29)	(-5.54)	(2.03)	(-1.58)	(-1.58)	(-0.80)	(1.97)	(0.54)	(-0.02)	(1.70)	(-2.24)	(2.68)
11. Red meat	0.49*	-0.36*	-0.35*	0.56*	0.11	-0.06	0.12	-0.30*	-0.11*	0.28	-1.52*	-0.35*	0.25*	-0.04	-0.72*	-0.22*	0.08*	-0.35*	-0.03*	2.22*
	(2.80)	(-2.37)	(-3.39)	(2.37)	(0.86)	(-1.42)	(1.45)	(-2.65)	(-2.50)	(1.59)	(-7.27)	(-2.02)	(2.94)	(-0.88)	(-4.85)	(-4.10)	(3.21)	(-4.13)	(-2.41)	(7.54)
12. Poultry	-0.56*	0.38	-0.09	-0.31	0.53*	-0.02	0.05	-0.87*	-0.11	-0.40*	-0.31	-2.45*	-0.39*	-0.01	-0.10	0.18*	-0.01	0.36*	0.02*	1.94*
	(-3.23)	(0.97)	(-1.07)	(-1.70)	(2.05)	(-0.45)	(0.22)	(-3.62)	(-1.73)	(-2.82)	(-1.77)	(-5.41)	(-3.65)	(-0.33)	(-1.10)	(2.85)	(-0.77)	(3.26)	(2.50)	(4.70)
13. Eggs	-0.49*	0.22	0.07	0.20	-0.43	-0.19	1.22*	-0.96*	-0.48*	-0.55*	1.13*	-2.14*	-0.67*	0.33	0.06	-0.44*	0.07	0.40*	0.04	1.93*
	(-2.09)	(1.36)	(0.40)	(0.56)	(-1.74)	(-1.28)	(3.49)	(-3.44)	(-2.81)	(-3.06)	(2.99)	(-4.45)	(-2.56)	(1.36)	(0.25)	(-2.43)	(1.08)	(2.16)	(1.06)	(5.05)
14. Fish†	-0.90*	0.52*	0.00	0.11	-0.05	-0.01	0.74*	0.16	0.30*	-0.11	0.02	0.04	0.11	-1.59*	0.07	0.08	-0.04	-0.10	-0.05*	0.80*
	(-3.12)	(2.52)	(0.08)	(0.88)	(-0.65)	(-0.51)	(3.55)	(1.95)	(3.28)	(-0.80)	(0.22)	(0.67)	(1.64)	(-10.24)	(1.02)	(1.56)	(-1.90)	(-1.54)	(-2.88)	(4.58)
15. Dairy	0.58*	-0.23	-0.09	0.56*	0.39*	0.04	0.07	-0.35*	-0.14*	0.30*	-1.07*	-0.16	0.02	0.05	-1.23*	-0.16*	0.02	-0.01	-0.02*	1.43*
	(3.94)	(-1.82)	(-1.56)	(3.14)	(2.66)	(1.45)	(0.98)	(-3.41)	(-2.92)	(2.74)	(-6.14)	(-1.28)	(0.14)	(0.45)	(-6.76)	(-3.94)	(1.16)	(-0.17)	(-2.78)	(6.99)
16. Fats†	-0.17	0.04	0.25	0.29	0.12	0.22*	-0.31	0.19	-0.14	0.05	-0.48*	0.62*	-0.23*	0.12	-0.25 *	-0.55	0.26*	-0.03	0.11*	1.07*
	(-1.22)	(0.45)	(1.68)	(1.52)	(1.18)	(2.13)	(-1.85)	(1.36)	(-1.69)	(0.17)	(-2.51)	(2.29)	(-1.96)	(0.97)	(-2.26)	(-0.66)	(2.86)	(-0.43)	(2.32)	(2.88)
17. Coffee†	0.21	0.87*	0.42*	-0.10	-0.69*	-0.23*	0.45*	0.12	0.52*	-0.03	0.62*	-0.07	0.11	-0.24*	0.10	0.78*	-0.97*	0.48*	-0.01	1.19*
	(1.29)	(5.08)	(2.76)	(-1.20)	(-3.95)	(-2.14)	(2.79)	(0.99)	(4.60)	(-0.17)	(3.73)	(-0.75)	(1.10)	(-2.38)	(1.26)	(4.97)	(-13.15)	(4.35)	(-0.49)	(5.97)
18. Soft drink†	0.04	-0.02	-0.94*	-0.59*	-0.20	0.23*	-0.48*	0.09	0.31*	0.32 -1.08*	1.28*	0.26*	-0.26	-0.03	-0.06	0.19*	-1.27*	0.01	1.72*	
	(-0.17)	(0.10)	(-3.71)	(-2.25)	(-0.73)	(2.82)	(-2.61)	(-0.15)	(2.69)	(1.66)	(-3.72)	(3.26)	(2.00)	(-1.71)	(-0.21)	(-0.60)	(3.41)	(-5.79)	(0.38)	(4.16)
19. Other food	-0.37*	0.43*	-0.50*	0.19*	-0.08	0.06	-0.17	0.55	-0.32*	-0.37*	-0.19	0.37	0.10	-0.39*	-0.13*	0.46*	-0.01	0.04	-0.95*	0.67*
	(-2.60)	(3.85)	(-3.17)	(2.03)	(-0.87)	(0.76)	(-1.51)	(5.55)	(-3.64)	(-3.44)	(-1.87)	(3.47)	(1.39)	(-3.91)	(-2.50)	(4.43)	(-0.36)	(0.65)	(-10.53)	(3.66)

Notes: This table shows consumption quantity elasticities with respect to food prices and total expenditure with t-value at median in parenthesis (* $p < 0.05$). We estimate these elasticities using the whole sample with the cross-sectional model. Refer to the following full name of each food category for row title with †; 1. Rice 2. Maize 3. Wheat and other cereals 4. Cassava 5. Roots, Tubers and, other starches 6. Sugar 7. Pulses 8. Nuts and seeds 9. Vegetables 10. Fruit 11. Red meat 12. Poultry 13. Eggs 14. Fish and seafood 15. Dairy 16. Fats and oils 17. Coffee, tea, and cocoa 18. Soft drink and juice 19. Other food Expenditure

Table 3.3: Price and Total Expenditure Elasticities for the Whole Sample: Panel Model

	Food Category																			Expenditure
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. Rice	-2.38*	0.57	-0.12	0.43	-0.30	-0.03	0.36	-0.28	0.13	0.01	0.34*	-0.33	0.00	-0.39*	0.26*	-0.19*	0.05	-0.05	-0.03*	1.30*
	(-5.16)	(2.08)	(-1.27)	(1.79)	(-1.54)	(-0.70)	(2.32)	(-1.92)	(1.75)	(-0.12)	(2.04)	(-1.72)	(0.21)	(-2.53)	(1.98)	(-2.14)	(1.77)	(-0.77)	(-2.06)	(4.46)
2. Maize	0.35	-1.06	0.34	-0.22	0.10	0.17	-0.04	-0.48	-0.07	0.24	-0.03	0.08	0.02	0.09	0.05	0.08	0.07	0.07*	0.02*	0.68
	(1.67)	(-3.28)	(1.89)	(-1.10)	(0.65)	(2.39)	(-0.60)	(-1.92)	(-1.34)	(1.30)	(-0.24)	(0.64)	(0.61)	(1.04)	(1.21)	(1.70)	(2.24)	(1.27)	(2.19)	(1.19)
3. Wheat†	-0.18	0.86	-2.76*	0.34	0.86	-0.12	-0.00	0.51	0.03	-0.02	-0.58*	-0.01	0.00	0.16	-0.12	-0.01	0.07*	-0.36*	-0.04*	1.36*
	(-1.34)	(2.67)	(-6.06)	(1.73)	(3.03)	(-1.64)	(0.03)	(2.32)	(0.17)	(-0.17)	(-2.93)	(-0.26)	(0.03)	(1.54)	(-1.57)	(-0.37)	(2.14)	(-2.72)	(-2.08)	(4.95)
4. Cassava	0.55*	-0.36	0.30*	-1.49*	-0.31	-0.01	-0.29*	-0.08	0.02	-0.01	0.70*	0.08	0.10	0.18	0.34*	0.10	-0.02	-0.22	0.03	0.18
	(2.17)	(-1.14)	(2.03)	(-4.27)	(-1.23)	(-0.30)	(-2.14)	(-0.21)	(0.28)	(0.20)	(2.28)	(0.22)	(1.38)	(1.44)	(1.96)	(1.28)	(-1.14)	(-1.91)	(1.85)	(1.16)
5. Roots†	-0.33	0.15	0.64*	-0.36	-2.38*	-0.08	-0.79*	0.28	0.27*	-0.12	0.21	0.28	-0.17	-0.19	0.40*	0.06	-0.06*	-0.04	-0.01	1.24*
	(-1.35)	(0.46)	(2.38)	(-1.40)	(-4.32)	(-1.41)	(-2.57)	(1.33)	(2.34)	(-0.74)	(1.27)	(0.73)	(-1.55)	(-1.38)	(2.26)	(1.08)	(-1.98)	(-0.43)	(-0.99)	(2.65)
6. Sugar	-0.08	0.77*	-0.19	-0.08	-0.18	-0.72*	0.38*	-0.02	0.14	-0.10	-0.04	-0.20	-0.15	-0.01	0.15*	0.15*	-0.07	0.15*	-0.02	1.05*
	(-0.59)	(4.33)	(-1.88)	(-1.07)	(-1.58)	(-4.35)	(2.65)	(-0.18)	(1.81)	(-1.23)	(-0.25)	(-1.43)	(-1.87)	(-0.19)	(2.31)	(2.01)	(-1.74)	(2.22)	(-1.04)	(6.16)
7. Pulses	0.52*	-0.10	0.01	-0.41*	-1.01*	0.21*	-1.55*	-0.05	0.16*	-0.06	0.26*	-0.04	0.34*	0.53*	0.19	-0.08	0.11*	-0.17	-0.02	0.75*
	(2.63)	(-0.66)	(0.37)	(-2.42)	(-3.23)	(2.51)	(-5.64)	(-0.30)	(2.07)	(-0.25)	(2.10)	(-0.02)	(2.84)	(3.09)	(1.79)	(-1.20)	(2.45)	(-1.79)	(-1.32)	(3.03)
8. Nuts†	-0.55	-1.92*	0.68	-0.19	0.52	-0.02	-0.11	-1.92*	-0.03	0.47	-0.51	-1.10	-0.38	0.24	-0.33	0.06	0.00	-0.06	0.08*	1.41*
	(-1.56)	(-2.21)	(1.88)	(-0.72)	(1.22)	(-0.26)	(-0.74)	(-3.76)	(-0.42)	(1.37)	(-1.55)	(-1.62)	(-1.87)	(1.17)	(-1.54)	(0.64)	(0.09)	(-0.35)	(2.23)	(2.67)
9. Vegetables	0.23*	-0.16*	0.07	-0.03	0.36*	0.08*	0.15*	0.01	-0.65*	-0.23*	0.09	0.03	-0.08*	0.22*	-0.11*	-0.03	0.07*	0.15*	-0.04*	0.64*
	(3.46)	(-2.00)	(1.45)	(0.10)	(3.91)	(2.14)	(2.38)	(0.37)	(-10.41)	(-3.02)	(1.29)	(0.79)	(-2.16)	(4.83)	(-2.92)	(-1.05)	(4.41)	(3.41)	(-3.45)	(5.22)
10.Fruit	0.04	0.77	-0.00	-0.05	-0.16	-0.06	-0.08	0.43	-0.31*	-1.06*	0.50	-0.27	-0.03	0.20	0.21	-0.03	-0.02	0.10	-0.02	0.92*
	(0.15)	(1.32)	(-0.05)	(-0.18)	(-0.66)	(-0.96)	(-0.57)	(1.30)	(-2.01)	(-3.63)	(1.57)	(-0.75)	(-0.40)	(1.05)	(1.47)	(-0.30)	(-0.56)	(0.93)	(-1.12)	(2.29)
11. Red meat	0.31	-0.23	-0.48*	0.61*	0.17	-0.04	0.14	-0.30*	-0.01	0.28	-1.54*	-0.54*	0.27*	-0.11	-0.82*	-0.21*	0.08*	-0.30*	-0.04*	2.08*
	(1.50)	(-1.27)	(-3.66)	(2.70)	(1.11)	(-0.83)	(1.20)	(-2.32)	(-0.19)	(1.52)	(-6.53)	(-2.45)	(3.12)	(-1.33)	(-4.73)	(-3.83)	(3.23)	(-3.08)	(-3.10)	(7.19)
12. Poultry	-0.37*	-0.12	-0.03	0.01	0.19	-0.10	-0.11	-0.55*	-0.05	-0.19	-0.49	-1.83*	-0.30*	0.13	-0.19	0.09	-0.04	0.20	0.02	2.15*
	(-2.08)	(-0.64)	(-0.54)	(-0.11)	(0.20)	(-1.45)	(-0.89)	(-2.27)	(-0.79)	(-1.33)	(-1.94)	(-3.76)	(-2.48)	(0.82)	(-1.42)	(1.13)	(-1.41)	(1.53)	(1.41)	(3.55)
13. Eggs	-0.04	0.02	-0.02	0.40	-0.84	-0.32	1.24*	-1.00*	-0.38*	-0.14	1.22*	-1.67*	-0.82*	0.33	0.08	-0.50*	0.08	0.14	-0.03	2.16*
	(-0.15)	(0.25)	(-0.11)	(1.00)	(-1.88)	(-1.75)	(2.89)	(-2.70)	(-2.18)	(-0.94)	(2.66)	(-3.18)	(-2.54)	(1.26)	(0.24)	(-2.21)	(1.07)	(0.60)	(-0.68)	(4.25)
14. Fish†	-0.56*	0.21	0.19	0.21	-0.23	0.00	0.58*	0.19	0.26*	0.18	-0.08	0.27	0.11	-1.34*	0.05	0.00	-0.02	-0.18	-0.05*	0.90*
	(-2.33)	(1.16)	(1.74)	(1.19)	(-1.48)	(-0.08)	(2.76)	(1.57)	(2.70)	(1.40)	(-0.66)	(1.46)	(1.62)	(-8.78)	(0.87)	(0.07)	(-0.95)	(-1.97)	(-2.61)	(4.14)
15. Dairy	0.41*	0.07	-0.15*	0.45*	0.62*	0.08*	0.18	-0.29*	-0.19*	0.19	-1.23*	-0.32*	0.03	0.02	-1.08*	-0.11*	0.02	0.01	-0.03*	1.50*
	(2.21)	(0.64)	(-2.29)	(2.43)	(3.38)	(2.10)	(1.76)	(-2.60)	(-3.66)	(1.30)	(-6.02)	(-2.02)	(0.13)	(0.21)	(-5.51)	(-2.52)	(1.13)	(-0.18)	(-3.45)	(7.40)
16. Fats†	-0.48	0.35	-0.01	0.19	0.18	0.16	-0.17	0.11	-0.11	-0.04	-0.44*	0.37	-0.26	-0.01	-0.17	-0.54	0.28*	-0.02	0.08	0.96*
	(-1.71)	(1.58)	(-0.10)	(1.07)	(1.16)	(1.54)	(-1.08)	(0.85)	(-1.22)	(-0.25)	(-2.04)	(1.47)	(-1.74)	(-0.08)	(-1.56)	(-0.82)	(2.36)	(-0.11)	(1.80)	(2.26)
17. Coffee†	0.38*	0.94*	0.40*	-0.22	-0.49*	-0.22	0.62*	0.01	0.43*	-0.10	0.61*	-0.31	0.14	-0.12	0.10	0.85*	-0.96*	0.39*	0.07	1.18*
	(1.96)	(3.86)	(2.40)	(-1.68)	(-2.59)	(-1.79)	(2.64)	(0.09)	(3.23)	(-0.67)	(3.10)	(-1.47)	(1.15)	(-1.14)	(1.23)	(4.05)	(-11.92)	(3.09)	(1.90)	(4.64)
18. Soft drink†	-0.22	0.34	-0.88*	-0.74*	-0.16	0.17	-0.46	-0.11	0.35*	0.17	-0.95*	0.73	0.09	-0.45*	0.00	-0.05	0.15*	-1.09*	0.03	1.75*
	(-0.82)	(0.91)	(-2.49)	(-2.04)	(-0.57)	(1.79)	(-1.89)	(-0.60)	(2.25)	(0.76)	(-2.44)	(1.55)	(0.63)	(-2.05)	(-0.12)	(-0.38)	(2.34)	(-3.77)	(1.17)	(3.33)
19. Other food	-0.30	0.48*	-0.26*	0.26*	-0.11	-0.08	-0.17	0.51*	-0.31*	-0.12	-0.36*	0.31*	-0.05	-0.35*	-0.20*	0.34*	0.10*	0.11	-0.92*	0.62*
	(-1.93)	(3.77)	(-2.08)	(2.31)	(-0.72)	(-0.93)	(-1.34)	(4.69)	(-3.14)	(-0.95)	(-2.58)	(2.48)	(-0.31)	(-3.39)	(-2.82)	(3.27)	(2.01)	(1.66)	(-10.67)	(2.98)

Notes: This table shows consumption quantity elasticities with respect to food prices and total expenditure with t-value at median in parenthesis (* $p < 0.05$). We estimate these elasticities using the whole sample with the panel model. Refer to the following full name of each food category for row title with †; 1. Rice 2. Maize 3. Wheat and other cereals 4. Cassava 5. Roots, Tubers and, other starches 6. Sugar 7. Pulses 8. Nuts and seeds 9. Vegetables 10. Fruit 11. Red meat 12. Poultry 13. Eggs 14. Fish and seafood 15. Dairy 16. Fats and oils 17. Coffee, tea, and cocoa 18. Soft drink and juice 19. Other food Expenditure

Table 3.4: Own-Price Elasticities by Quartile: Panel Model

Food Category	Per-capita Total Expenditure Quartile			
	Q1	Q2	Q3	Q4
1. Rice	-3.55*	-2.49*	-1.98*	-1.37*
	(-5.02)	(-5.43)	(-5.43)	(-4.61)
2. Maize	-1.09*	-1.08*	-1.05*	-0.91
	(-4.01)	(-3.63)	(-2.98)	(-1.81)
3. Wheat and other cereals	-4.05*	-2.92*	-2.19*	-1.32*
	(-6.27)	(-6.39)	(-6.04)	(-4.84)
4. Cassava	-1.39*	-1.46*	-1.53*	-1.66*
	(-4.61)	(-4.69)	(-4.29)	(-3.62)
5. Roots, Tubers, and other starches	-2.74*	-2.43*	-2.26*	-2.06*
	(-4.18)	(-4.70)	(-4.56)	(-3.69)
6. Sugar	-0.89*	-0.74*	-0.64*	-0.50*
	(-5.00)	(-5.07)	(-4.23)	(-2.05)
7. Pulses	-1.71*	-1.58*	-1.48*	-1.27*
	(-6.20)	(-6.17)	(-5.59)	(-3.93)
8. Nuts and seeds	-2.41*	-1.99*	-1.67*	-1.23*
	(-3.88)	(-3.93)	(-3.71)	(-2.95)
9. Vegetables	-0.60*	-0.64*	-0.68*	-0.77*
	(-9.60)	(-11.26)	(-11.21)	(-9.46)
10. Fruit	-1.58*	-1.10*	-0.83*	-0.52
	(-4.15)	(-4.07)	(-3.02)	(-1.62)
11. Red meat	-1.67*	-1.56*	-1.50*	-1.43*
	(-4.93)	(-7.07)	(-7.80)	(-6.97)
12. Poultry	-3.10*	-1.94*	-1.28*	-0.46*
	(-4.02)	(-4.02)	(-3.48)	(-2.13)
13. Eggs	-0.94	-0.83*	-0.77*	-0.71*
	(-1.94)	(-2.81)	(-2.85)	(-2.87)
14. Fish and seafood	-1.52*	-1.36*	-1.27*	-1.15*
	(-8.29)	(-9.45)	(-9.12)	(-8.31)
15. Dairy	-0.91*	-1.07*	-1.15*	-1.26*
	(-3.46)	(-5.92)	(-6.35)	(-5.78)
16. Fats and oils	-0.76	-0.57	-0.42	-0.20
	(-1.63)	(-0.84)	(-0.41)	(0.29)
17. Coffee, tea, and cocoa	-0.74*	-0.94*	-1.05*	-1.17*
	(-8.09)	(-13.89)	(-14.20)	(-11.49)
18. Soft drink and juice	-1.13*	-1.10*	-1.08*	-1.06*
	(-3.18)	(-4.16)	(-4.27)	(-3.70)
19. Other food	-0.88*	-0.91*	-0.95*	-1.01*
	(-9.99)	(-12.10)	(-11.71)	(-8.82)

Notes: This table shows own-price elasticities by per-capita total expenditure quartile with t-value at median in parenthesis (* $p < 0.05$). We estimate these elasticities with the panel model.

Table 3.5: Nutrient Elasticities for the Whole Sample

	Food Category																			Expenditure
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
kcal	-0.09	-0.37	-0.02	-0.14	-0.09	0.05	-0.11	-0.24	-0.01	0.08	-0.04	-0.05	0.02	0.03	0.05	-0.01	0.05*	-0.04	0.00	0.93*
	(-0.18)	(-1.90)	(0.39)	(-0.79)	(-0.85)	(1.23)	(-2.05)	(-1.88)	(-0.81)	(1.02)	(-0.43)	(-0.33)	(0.62)	(0.73)	(1.43)	(0.14)	(2.78)	(-0.60)	(0.84)	(2.69)
protein	-0.04	-0.37*	-0.02	-0.04	-0.10	0.08	-0.18*	-0.27*	0.01	0.12	-0.17*	-0.13	0.06	-0.02	-0.05	-0.01	0.05*	-0.06	-0.01	1.04*
	(0.12)	(-2.27)	(0.72)	(-0.24)	(-1.14)	(1.76)	(-2.50)	(-2.23)	(-0.07)	(1.42)	(-2.31)	(-0.76)	(1.68)	(0.22)	(-0.37)	(-0.14)	(2.83)	(-1.39)	(-0.92)	(3.45)
fat	-0.10	-0.26	-0.06	0.09	0.21	0.07	-0.07	-0.30	-0.08*	0.07	-0.43*	-0.14	-0.08	0.04	-0.23*	-0.15	0.11*	-0.04	0.02	1.17*
	(-0.86)	(-1.18)	(-0.01)	(0.68)	(1.47)	(1.21)	(-1.33)	(-1.89)	(-2.04)	(0.94)	(-3.63)	(-0.29)	(-1.79)	(0.85)	(-2.97)	(-0.80)	(2.41)	(-0.67)	(1.84)	(3.10)
sugar	0.00	0.46*	-0.22	-0.13	-0.25*	-0.37*	0.06	0.00	0.06	-0.11	-0.07	-0.11	-0.06	0.04	0.07*	0.07	-0.02	0.02	-0.02	1.06*
	(0.22)	(3.49)	(-2.33)	(-1.35)	(-2.16)	(-3.71)	(0.43)	(-0.06)	(1.24)	(-1.24)	(-0.43)	(-0.95)	(-1.17)	(0.97)	(2.04)	(1.60)	(-0.97)	(0.44)	(-1.43)	(5.40)
iron	0.11	-0.55*	0.08	-0.20	-0.13	0.10	-0.23*	-0.28	-0.04	0.09	0.02	-0.02	0.04	0.10	0.06	0.02	0.06*	-0.01	0.00	0.84*
	(0.72)	(-2.63)	(1.09)	(-1.18)	(-1.05)	(1.99)	(-3.24)	(-1.92)	(-1.53)	(1.05)	(0.39)	(-0.30)	(1.20)	(1.46)	(1.74)	(0.86)	(2.78)	(-0.26)	(0.21)	(2.49)
zinc	-0.06	-0.49*	0.03	-0.07	-0.06	0.09	-0.13*	-0.31*	-0.03	0.12	-0.11	-0.11	0.06	0.04	-0.01	-0.01	0.06*	-0.03	0.00	0.97*
	(0.07)	(-2.49)	(0.82)	(-0.41)	(-0.71)	(1.94)	(-2.18)	(-2.27)	(-1.16)	(1.34)	(-1.18)	(-0.65)	(1.54)	(0.81)	(0.53)	(-0.05)	(3.02)	(-0.70)	(0.10)	(2.95)
vitamin A	-0.22	0.15	0.27	-0.06	-0.84*	0.03	-0.33	0.14	-0.07	-0.12	-0.06	0.22	-0.16*	-0.03	0.08	-0.15	0.07*	0.01	0.00	0.98*
	(-1.40)	(1.16)	(1.89)	(-0.40)	(-3.56)	(1.08)	(-2.06)	(1.37)	(-1.35)	(-1.11)	(-0.83)	(1.19)	(-2.17)	(-0.07)	(0.87)	(-0.85)	(2.19)	(0.08)	(0.43)	(3.41)
total folate	0.18	-0.32*	0.12	-0.28	-0.42*	0.10*	-0.57*	-0.18	-0.03	-0.02	0.10	-0.05	0.08	0.20*	0.09	-0.01	0.06*	-0.05	-0.01	0.84*
	(1.24)	(-2.24)	(1.76)	(-1.89)	(-2.65)	(2.32)	(-5.13)	(-1.74)	(-0.82)	(0.28)	(1.57)	(-0.44)	(1.50)	(2.63)	(1.90)	(-0.02)	(2.84)	(-1.29)	(-0.95)	(3.16)

Notes: This table shows nutrient elasticities with respect to food prices and total expenditure with t-value at median in parenthesis (* $p < 0.05$). 1. Rice 2. Maize 3. Wheat and other cereals 4. Cassava 5. Roots, Tubers and, other starches 6. Sugar 7. Pulses 8. Nuts and seeds 9. Vegetables 10. Fruit 11. Red meat 12. Poultry 13. Eggs 14. Fish and seafood 15. Dairy 16. Fats and oils 17. Coffee, tea, and cocoa 18. Soft drink and juice 19. Other food Expenditure

Bibliography

Abdulai, Awudu and Dominique Aubert, "A Cross-section Analysis of Household Demand for Food and Nutrients in Tanzania," *Agricultural Economics*, 2004, *31* (1), 67–79.

Agbola, Frank W, "Estimation of Food Demand Patterns in South Africa based on a Survey of Households," *Journal of Agricultural and Applied Economics*, 2003, *35* (3), 663–670.

Albouy, David, "The Unequal Geographic Burden of Federal Taxation," *Journal of Political Economy*, 2009, *117* (4), 635–667.

Alinovi, Luca, Marco D'errico, Erdgin Mane, and Donato Romano, "Livelihoods Strategies and Household Resilience to Food Insecurity: An Empirical Analysis to Kenya," in "conference organized by the European Report of Development, Dakar, Senegal, june" 2010, pp. 28–30.

Almås, Ingvild and Ashild Johnsen, "The Cost-of-living in China: Implications for Inequality and Poverty," Technical Report, Institute for Samfunnsokonomi, Department of Economics 2012.

Andersen, Steffen, Glenn W Harrison, Morten Igel Lau, and Elisabet Rutström, "Elicitation Using Multiple Price List Formats," *Experimental Economics*, 2006, *9* (4), 383–405

Anderson, Simon P, Andre De Palma, and Jacques François Thisse, *Discrete Choice Theory of Product Differentiation*, MIT press, 1992.

Atkin, David, "Trade, Tastes, and Nutrition in India," *American Economic Review*, 2013, *103* (5), 1629–63.

Babcock, Bruce A, "Using Cumulative Prospect Theory to Explain Anomalous Crop Insurance Coverage Choice," *American Journal of Agricultural Economics*, 2015, *97* (5), 1371–1384.

Banks, James, Richard Blundell, and Arthur Lewbel, "Quadratic Engel Curves and Consumer Demand," *Review of Economics and Statistics*, 1997, *79* (4), 527–539.

Barberis, Nicholas C, "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 2013, *27* (1), 173–96.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C Teitelbaum, "The Nature of Risk Preferences: Evidence from Insurance Choices," *American Economic Review*, 2013, *103* (6), 2499–2529.

Beatty, Timothy KM, "Do the Poor Pay More for Food? Evidence from the United Kingdom," *American Journal of Agricultural Economics*, 2010, *92* (3), 608–621.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak, "Measuring Utility by a Single-response Sequential Method," *Systems Research and Behavioral Science*, 1964, *9* (3), 226–232.

Bils, Mark, "Measuring the Growth from Better and Better goods," Technical Report, National Bureau of Economic Research 2004.

___ **and Peter J Klenow**, "Quantifying Quality Growth," *American Economic Review*, 2001, *91* (4), 1006–1030.

Binswanger, Hans P, "Attitudes Toward Risk: Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, 1980, *62* (3), 395–407.

Binswanger-Mkhize, Hans P, "Is There Too Much Hype About Index-based Agricultural Insurance?," *Journal of Development Studies*, 2012, *48* (2), 187–200.

Bocquého, Géraldine, Florence Jacquet, and Arnaud Reynaud, "Expected Utility or Prospect Theory Maximisers? Assessing Farmers' Risk Behaviour from Field-experiment Data," *European Review of Agricultural Economics*, 2013, *41* (1), 135–172.

Bopape, Lesiba and Robert Myers, "Analysis of Household Demand for Food in South Africa: Model Selection, Expenditure Endogeneity, and the Influence of Socio-demographic Effects," in "Proceedings of the African Econometrics Society Annual Conference" Citeseer 2007, pp. 4–6.

Brandt, Loren and Carsten A Holz, "Spatial Price Differences in China: Estimates and Implications," *Economic Development and Cultural Change*, 2006, *55* (1), 43–86.

Broda, Christian and David E Weinstein, "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, 2010, *100* (3), 691–723.

Brown, Jeffrey R, Jeffrey R Kling, Sendhil Mullainathan, and Marian V Wrobel, "Why Don't People Insure Late-life Consumption? A Framing Explanation of the Under-annuitization Puzzle," *American Economic Review*, 2008, *98* (2), 304–09.

Carter, Michael, Alain de Janvry, Elisabeth Sadoulet, Alexander Sarris et al., "Index-based Weather Insurance for Developing Countries: A Review of Evidence and a Set of Propositions for Up-scaling," *Development Policies working paper*, 2014, *111.*

Casaburi, Lorenzo and Jack Willis, "Time vs. State in Insurance: Experimental Evidence from Contract Farming in Kenya," *American Economic Review*, 2018, p. forthcoming.

Christensen, Laurits R, Dale W Jorgenson, and Lawrence J Lau, "Transcendental Logarithmic Utility Functions," *The American Economic Review*, 1975, *65* (3), 367–383.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery, "Barriers to Household Risk Management: Evidence from India," *American Economic Journal: Applied Economics*, 2013, *5* (1), 104–35.

Cox, Thomas L and Michael K Wohlgenant, "Prices and Quality Effects in Cross-sectional Demand Analysis," *American Journal of Agricultural Economics*, 1986, *68* (4), 908–919.

Deaton, Angus, "Quality, Quantity, and Spatial Variation of Price," *The American Economic Review*, 1988, pp. 418–430.

____ **and John Muellbauer**, "An Almost Ideal Demand System," *The American Economic Review*, 1980, *70* (3), 312–326.

____ **and Olivier Dupriez**, "Spatial Price Differences within Large Countries," Technical Report, Princeton University 2011. Manuscript, Princeton University, July.

Dercon, Stefan, "Growth and Shocks: Evidence from Rural Ethiopia," *Journal of Development Economics*, 2004, *74* (2), 309–329.

____ **and Luc Christiaensen**, "Consumption Risk, Technology Adoption and Poverty Traps: Evidence from Ethiopia," *Journal of Development Economics*, 2011, *96* (2), 159–173.

____ **, Ruth Vargas Hill, Daniel Clarke, Ingo Outes-Leon, and Alemayehu Seyoum Taffesse**, "Offering Rainfall Insurance to Informal Insurance Groups: Evidence from a Field Experiment in Ethiopia," *Journal of Development Economics*, 2014, *106*, 132–143.

Diewert, W Erwin, "Axiomatic and Economic Approaches to International Comparisons," in "International and Interarea Comparisons of Income, Output, and Prices," University of Chicago Press, 1999, pp. 13–107.

Dong, Diansheng, Brian W Gould, and Harry M Kaiser, "Food Demand in Mexico: an Application of the Amemiya-Tobin Approach to the Estimation of a Censored Food System," *American Journal of Agricultural Economics*, 2004, *86* (4), 1094–1107.

Du, Xiaodong, Hongli Feng, and David A Hennessy, "Rationality of Choices in Subsidized Crop Insurance Markets," *American Journal of Agricultural Economics*, 2016, *99* (3), 732–756.

Ecker, Olivier and Matin Qaim, "Analyzing Nutritional Impacts of Policies: An Empirical Study for Malawi," *World Development*, 2011, *39* (3), 412–428.

Feenstra, Robert C, "New Product Varieties and the Measurement of International Prices," *The American Economic Review*, 1994, pp. 157–177.

____, Mingzhi Xu, and Alexis Antoniades, "What is the Price of Tea in China? Towards the Relative Cost of Living in Chinese and US Cities," Technical Report, National Bureau of Economic Research 2017.

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, and Andrew Hoell, "The Climate Hazards Infrared Precipitation with Stations—A New Environmental Record for Monitoring Extremes," *Scientific Data*, 2015, *2*, 150066.

Gäl, Aykut, Cuma Akbay, Celile Özcicek, Remziye Özel, and Aysel Ozdes Akbay, "Expenditure Pattern for Food Away from Home Consumption in Turkey," *Journal of International Food & Agribusiness Marketing*, 2007, *19* (4), 31–43.

Gaurav, Sarthak, Shawn Cole, and Jeremy Tobacman, "Marketing Complex Financial Products in Emerging Markets: Evidence from Rainfall Insurance in India," *Journal of Marketing Research*, 2011, *48* (SPL), S150–S162.

Giné, Xavier and Dean Yang, "Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi," *Journal of Development Economics*, 2009, *89* (1), 1–11.

___ , **Dean Karlan, and Muthoni Ngatia**, "Social Networks, Financial Literacy and Index Insurance," Technical Report, World Bank, Washington, DC 2013.

___ , **Robert Townsend, and James Vickery**, "Patterns of Rainfall Insurance Participation in Rural India," *The World Bank Economic Review*, 2008, *22* (3), 539–566.

Grace, Martin F, Robert W Klein, Paul R Kleindorfer, and Michael R Murray, *Catastrophe Insurance: Consumer Demand, Markets and Regulation*, Vol. 1, Springer, 2003.

GS1, "GTIN Allocation Rules," Technical Report, GS1,Brussels, Belgium 2007. Based on the GS1 General Specifications Section 2.1, Brussels, Belgium.

Handbury, Jessie and David E Weinstein, "Goods Prices and Availability in Cities," *The Review of Economic Studies*, 2015, *82* (1), 258–296.

Hausman, Jerry, Gregory Leonard, and J Douglas Zona, "Competitive Analysis with Differentiated Products," *Annales d'Economie et de Statistique*, 1994, pp. 159–180.

Hill, Robert J, "A Taxonomy of Multilateral Methods for Making International Comparisons of Prices and Quantities," *Review of Income and Wealth*, 1997, *43* (1), 49–69.

Hill, Ruth Vargas, John Hoddinott, and Neha Kumar, "Adoption of Weather Index Insurance: Learning from Willingness to Pay among a Panel of Households in Rural Ethiopia," *Agricultural Economics*, 2013, *44* (4-5), 385–398.

___ , **Miguel Robles, and Francisco Ceballos**, "Demand for a Simple Weather Insurance Product in India: Theory and Evidence," *American Journal of Agricultural Economics*, 2016, *98* (4), 1250–1270.

Huang, Kuo S, "Nutrient Elasticities in a Complete Food Demand System," *American Journal of Agricultural Economics*, 1996, *78* (1), 21–29.

___ **and Biing-Hwan Lin**, "Estimation of Food Demand and Nutrient Elasticities from Household Survey Data," Technical Report, U.S. Department of Agriculture 2000. Technical Bulletin No. 1887.

J-PAL, "Make it Rain," J-PAL Bulletin, Abdul Latif Jameel Poverty Action Lab, Center for Effective Global Action, and Agricultural Technology Adoption Initiative. 2016.

Johnson, Eric J, John Hershey, Jacqueline Meszaros, and Howard Kunreuther, "Framing, Probability Distortions, and Insurance Decisions," *Journal of Risk and Uncertainty*, 1993, *7* (1), 35–51.

Jolliffe, Dean, "Poverty, Prices, and Place: How Sensitive is the Spatial Distribution of Poverty to Cost of Living Adjustments?," *Economic Inquiry*, 2006, *44* (2), 296–310.

Kasteridis, Panagiotis, Steven T Yen, and Cheng Fang, "Bayesian Estimation of a Censored Linear Almost Ideal Demand System: Food Demand in Pakistan," *American Journal of Agricultural Economics*, 2011, *93* (5), 1374–1390.

Khandelwal, Amit, "The Long and Short (of) Quality Ladders," *The Review of Economic Studies*, 2010, *77* (4), 1450–1476.

Knittel, Christopher R and Konstantinos Metaxoglou, "Estimation of Random-coefficient Demand Models: Two Empiricists' Perspective," *Review of Economics and Statistics*, 2014, *96* (1), 34–59.

Kőszegi, Botond and Matthew Rabin, "A Model of Reference-dependent Preferences," *The Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

—— **and** —— , "Reference-dependent Risk Attitudes," *American Economic Review*, 2007, *97* (4), 1047–1073.

—— **and** —— , "Reference-dependent Consumption Plans," *American Economic Review*, 2009, *99* (3), 909–36.

Lazaro, Edith, Abdoul G Sam, and Stanley R Thompson, "Rice Demand in Tanzania: An Empirical Analysis," *Agricultural Economics*, 2017, *48* (2), 187–196.

Lewbel, Arthur and Krishna Pendakur, "Tricks with Hicks: The EASI Demand System," *American Economic Review*, 2009, *99* (3), 827–63.

Liu, Elaine M, "Time to Change What to Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China," *Review of Economics and Statistics*, 2013, *95* (4), 1386–1403.

Liu, Yanyan, Kevin Chen, Ruth Vargas Hill, and Chengwei Xiao, "Delayed Premium Payment, Insurance Adoption, and Household Investment in Rural China," *IFPRI Discussion Paper 01306*, 2017.

Ma, Hengyun, Jikun Huang, Frank Fuller, and Scott Rozelle, "Getting Rich and Eating Out: Consumption of Food Away from Home in Urban China," *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 2006, *54* (1), 101–119.

Mao, Chang Xuan, Robert Colwell, and Jing Chang, "Interpolating, Extrapolating, and Comparing Incidence-based Species Accumulation Curves," *Ecology*, 2004, *85* (10), 2717–2727.

—— **,** —— **, and** —— , "Estimating the Species Accumulation Curve using Mixtures," *Biometrics*, 2005, *61* (2), 433–441.

Meyerhoefer, Chad D, Christine K Ranney, and David E Sahn, "Consistent Estimation of Censored Demand Systems using Panel Data," *American Journal of Agricultural Economics*, 2005, *87* (3), 660–672.

Nevo, Aviv, "A Practitioner's Guide to Estimation of Random-coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 2000a, *9* (4), 513–548.

——, "Mergers with Differentiated Products: The Case of the Ready-to-eat Cereal Industry," *The RAND Journal of Economics*, 2000b, pp. 395–421.

——, "Measuring Market Power in the Ready-to-eat Cereal Industry," *Econometrica*, 2001, *69* (2), 307–342.

Newey, Whitney K, "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables," *Journal of Econometrics*, 1987, *36* (3), 231–250.

Novemsky, Nathan and Daniel Kahneman, "The Boundaries of Loss Aversion," *Journal of Marketing research*, 2005, *42* (2), 119–128.

Ochieng, Justus, Victor Afari-Sefa, Daniel Karanja, Radegunda Kessy, Srinivasulu Rajendran, and Silvest Samali, "How Promoting Consumption of Traditional African Vegetables Affects Household Nutrition Security in Tanzania," *Renewable Agriculture and Food Systems*, 2018, *33* (2), 105–115.

Perali, Federico and Jean-Paul Chavas, "Estimation of Censored Demand Equations from Large Cross-section Data," *American Journal of Agricultural Economics*, 2000, *82* (4), 1022–1037.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica*, 1998, pp. 497–527.

Ramakrishnan, Usha, "Prevalence of Micronutrient Malnutrition Worldwide," *Nutrition Reviews*, 2002, *60* (suppl_5), S46–S52.

Ruel, Marie T, Nicholas Minot, and Lisa Smith, *Patterns and Determinants of Fruit and Vegetable Consumption in Sub-Saharan Africa: A Multicountry Comparison*, WHO Geneva, 2005.

Sato, Kazuo, "The Ideal Log-change Index Number," *The Review of Economics and Statistics*, 1976, pp. 223–228.

Serfilippi, Elena, Michael Carter, and Catherine Guirkinger, "Insurance Contracts when Farmers "Greatly Value Certainty": Results from Field Experiments in Burkina Faso," 2016. Working paper, University of Namur (CRED).

Sheu, Gloria, "Price, Auality, and Variety: Measuring the Gains from Trade in Differentiated Products," *American Economic Journal: Applied Economics*, 2014, *6* (4), 66–89.

Sydnor, Justin, "(Over) Insuring Modest Risks," *American Economic Journal: Applied Economics*, 2010, *2* (4), 177–99.

Tanaka, Tomomi, Colin F Camerer, and Quang Nguyen, "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review*, 2010, *100* (1), 557–71.

Tanzania National Bureau of Statistics, "Basic Information Document-National Panel Survey (NPS 2008-2009)," Technical Report 2009.

Tonsor, Glynn T, "Producer Decision Making under Uncertainty: Role of Past Experiences and Question Framing," *American Journal of Agricultural Economics*, 2018, *100* (4), 1120–1135.

Tversky, Amos and Daniel Kahneman, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–292.

—— **and** —— , "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

Vartia, Yrjö O, "Ideal Log-change Index Numbers," *Scandinavian Journal of Statistics*, 1976, pp. 121–126.

Weinberger, Katinka and Ignas Swai, "Consumption of Traditional Vegetables in Central and Northeastern Tanzania," *Ecology of Food and Nutrition*, 2006, *45* (2), 87–103.

Weliwita, Ananda, David Nyange, and Hiroshi Tsujii, "Food Demand Patterns in Tanzania: A Censored Regression Analysis of Microdata," *Sri Lankan Journal of Agricultural Economics*, 2002, *5* (1381-2016-115751), 9.

Wooldridge, Jeffrey M, *Econometric Analysis of Cross Section and Panel data*, MIT press, 2010.

Zhen, Chen, Ian F Brissette, and Ryan Richard Ruff, "By Ounce or by Calorie: the Differential Effects of Alternative Sugar-sweetened Beverage Tax Strategies," *American Journal of Agricultural Economics*, 2014, *96* (4), 1070–1083.

Appendix A

Chapter 1 Appendix

I. Derivation of Partial Derivatives

Starting with equation (1.4) and (1.5), we take partial derivatives with respect to risk preferences.

1. Risk Aversion

$$\begin{aligned}\frac{\partial U_1}{\partial \sigma} =& \lambda\pi(p - pq_1)(\rho + L(1-\gamma))^{(1-\sigma)} \ln(\rho + L(1-\gamma)) + \lambda\pi(pq_1)(\rho + L)^{(1-\sigma)} \ln(\rho + L) \\ &+ \lambda\pi(1 - p - q_2 + pq_2)\rho^{(1-\sigma)} \ln\rho - \pi(q_2 - pq_2)(\gamma L - \rho)^{(1-\sigma)} \ln(\gamma L - \rho) \quad \lesseqgtr 0\end{aligned} \tag{A.1}$$

Because the first three terms are positive and the last term is negative, the sign of this partial derivative is ambiguous. If the absolute value of the last term of (1.4), utility gain from the outcome due to positive basis risk, is less than the first three terms of (1.4), utility loss from the outcomes in the loss domain, then it must be true that $\frac{\partial U}{\partial \sigma} > 0$. This can be the case when farmers do not consider the outcome from a positive basis risk event. Given that we only introduce a negative basis risk case (where farmers do not receive insurance payouts for their losses even when they purchased insurance) in the information session, we think our sample farmers do not take positive basis risk into account in insurance purchase decisions. Our empirical evidence of the positive effect of risk aversion supports this case.

$$\begin{aligned}\frac{\partial U_2}{\partial \sigma} =& \lambda\pi(p - pq_1)(L(1-\gamma))^{(1-\sigma)} \ln(L(1-\gamma)) + \lambda\pi(pq_1)L^{(1-\sigma)} \ln L \\ &- \pi(q_2 - pq_2)(\gamma L)^{(1-\sigma)} \ln(\gamma L) \quad \lesseqgtr 0\end{aligned} \tag{A.2}$$

Similarly, the sign of the partial derivative is ambiguous. If utility gain from a positive basis risk event is less than utility loss from the first two cases, then the marginal effect of risk aversion is positive. In other words, the more risk averse an individual is, the more likely the person purchases insurance.

2. Loss Aversion

$$\begin{aligned}\frac{\partial U_1}{\partial \lambda} =& -\pi(p-pq_1)(\rho+L(1-\gamma))^{(1-\sigma)} - \pi(pq_1)(\rho+L)^{(1-\sigma)} \\ & -\pi(1-p-q_2+pq_2)\rho^{(1-\sigma)} < 0\end{aligned} \tag{A.3a}$$

$$\frac{\partial U_2}{\partial \lambda} = -\pi(p-pq_1)(L(1-\gamma))^{(1-\sigma)} - \pi(pq_1)L^{(1-\sigma)} < 0 \tag{A.3b}$$

3. Nonlinear Probability Weighting

$$\begin{aligned}\frac{\partial U_1}{\partial \alpha} =& -\frac{\partial \pi(p-pq_1)}{\partial \alpha}\lambda(\rho+L(1-\gamma))^{(1-\sigma)} - \frac{\partial \pi(pq_1)}{\partial \alpha}\lambda(\rho+L)^{(1-\sigma)} \\ & -\frac{\partial \pi((1-p)(1-q_2)}{\partial \alpha}\lambda\rho^{(1-\sigma)} + \frac{\partial \pi(q_2-pq_2)}{\partial \alpha}(\gamma L-\rho)^{(1-\sigma)} \lesseqgtr 0\end{aligned} \tag{A.4a}$$

$$\begin{aligned}\frac{\partial U_2}{\partial \alpha} =& -\frac{\partial \pi(p-pq_1)}{\partial \alpha}\lambda(L(1-\gamma))^{(1-\sigma)} - \frac{\partial \pi(pq_1)}{\partial \alpha}\lambda L^{(1-\sigma)} \\ & +\frac{\partial \pi(q_2-pq_2)}{\partial \alpha}(\gamma L)^{(1-\sigma)} \lesseqgtr 0\end{aligned} \tag{A.4b}$$

$$\text{where} \quad \frac{\partial \pi(m)}{\partial \alpha} = -(-\ln m)^{\alpha}\ln(-\ln m)\exp\left(-(-\ln m)^{\alpha}\right)$$

$\frac{\partial \pi(m)}{\partial \alpha}$ is positive or negative, and this term is not monotonically increasing or decreasing in m. However, when probability m is smaller than about 0.1 the sign of this term is negative over $\alpha \in (0, 1.5]$, the range of non-linear probability weighting of the sample farmers. If the probability of drought (p) is large (which is what sample farmers report) and the probability of any basis risk (q_1 and q_2) is small enough to let $pq_1 < 0.1$ (the probability of the negative basis risk event) and $q_2 - pq_2 < 0.1$ (the probability of the positive basis risk event), then every term of (A.4a) except the second term is negative. $\frac{\partial U_1}{\partial \alpha} > 0$ only when the magnitude of the second term is greater than the magnitude of the other terms. Analogously, with a large probability of p (e.g., 0.4) and small probabilities of q_1 and q_2 (e.g., 0.17 and 0.15, respectively), the first and the third terms of (A.4b) are negative with the second term being positive. $\frac{\partial U_2}{\partial \alpha} > 0$ only when the magnitude of the second term is greater than the magnitude of the other two terms.

II. Experimental Auction Protocol

The auction was administered by the Game Master who administered the game. He led farmers and assistants though the various training rounds of the auction. The Assistants worked with the same two farmers as in the game, working through the various exercises in the auction, explaining farmers' outcomes, and recording bids into a worksheet. The following is the text used during enumerator training and auction administration:

This document is to be read by the Game Master (GM) in a way that is accessible to participants, except for sections that directly address Assistants.

1. Introduction

Now you will have the opportunity to buy the [WII sorghum/green gram but mainly, sorghum insurance] product that we introduced in the previous session. This policy has been designed by and will be sold by ACRE. Because this is an educational session to teach you about the benefits of insurance, we will provide varying levels of subsidies, determined by a lottery. Some of you will receive a higher subsidy than others, and some will not receive any subsidy it all depends on the lottery. Please remember that this subsidy will only apply to one seasons worth of insurance. If you buy insurance next season, it will only be offered at the market price.

This auction will be binding, but you will not have to pay today. You will sign an agreement to buy the amount of sorghum insurance you desire today, and a representative from ACRE will visit you in a couple days to collect payment. Therefore, take this auction seriously, but do not worry about paying today.

As in the previous session, you will work with an assistant in this session. At any point if you have questions, you should ask your assistant. We will show you one insurance product offered by ACRE that resembles the insurance product our team previously described to you.

The best strategy for this auction is to state the true amount of insurance you would purchase at each price. The auction does not work well if you state that you will only purchase insurance at very low prices to try to get a deal. The eventual prices of the insurance products are determined by a draw from an envelope, not by your bids. You will purchase insurance at the corresponding price drawn from the envelope if and only if you bid at least that much. You will not pay more or less than the price drawn from the envelope. Therefore, state exactly how much you are actually willingly to buy at each price.

This next point is very important: Nobody else's choices affect whether or not you will purchase insurance or the price you will pay for insurance. Only your own decisions and the prices drawn from the envelopes will affect whether or not you buy insurance, and how much insurance you buy.

This may seem complicated, but we will do a practice auction that should make things clear. Before we move on, your assistants will ask you about output and input values for sorghum (if you grow sorghum) or green gram (if you grow green gram ONLY) you expect this short rain season.

Has everyone finished? Great, lets move on to the first practice auction.

[Conduct Example auction using cookies: Farmers did a round where their bids can lead them to purchase real cookies with real money. We gave them small participation fees that could be used for this practice auction. The assistants asked how many cookies their farmers would like to buy at four different prices from Ksh 1 to Ksh 10. Then, we let one farmer draw a price card from an envelope to decide the revealed price. Farmers bought the number of cookies that they said they had wanted to purchase at the drawn price.]

2. Real Auction for WII

Now you are going to actually bid on WII insurance as you did for cookies before. Your bids are a commitment to pay real money for a real insurance policy. You will decide how many units of insurance you would like to purchase for your sorghum, given different prices. This may not be as easy as bidding on cookies, because you are more familiar with cookies than insurance. However, your assistants will help you to understand how this auction works. If you have any questions during the auction, feel free to ask your assistants. Remember that it is important you bid the true amount of insurance you would like to purchase at each price.

The coverage of 1 unit of insurance is Ksh 5,000. This means if you bought 1 unit of insurance, you can receive at maximum Ksh 5000 when rainfall is bad. If you bought 2 units of insurance, you can receive a maximum of Ksh 10,000 when rainfall is bad, and so on. However, you may not receive the full payout, or even receive no payout, even if you bought insurance. This payout depends on the overall rainfall for your squared area, not on your farm. Also remember that even when there is a drought, you can usually harvest something. Insurance is meant to make up the difference between what you expected to harvest and what you actually harvested. You will bid on the product to purchase and receive up to Ksh 500,000 worth of protection. This is 100 units. In addition, the price of 1 unit of insurance ranges from Ksh 50 to Ksh 1,000.

In the introduction of the auction, your assistants asked you output values for sorghum or green gram you expected to harvest, right? Based on the value, your assistants will help you to calculate how much it will cost to insure for sorghum at the different prices.

ASSISTANTS: Start from "Expected net value of sorghum produced". Suppose the value is Ksh 28,300. This value is between 25,000 (5 units of insurance) and 30,000 (6 units of insurance). Then, by referring the price table, start the conversation like: "[Farmer name], you said your net production value from sorghum is Ksh 28,300, right? Ok, lets start with the first price. The price of 1 unit of insurance is Ksh 50. If the maximum amount of insured is Ksh 25000, which is close to your harvest value, Ksh 28300, you need to buy 5 units of insurance, right? Then, you need to pay Ksh 250 because you buy 5 units with Ksh 50 per unit. Are you comfortable to pay Ksh 250 to be insured at maximum Ksh 25000? [In this moment, ASSISTANT should emphasize that Ksh 25000 is the maximum amount that the farmer can receive when bad rainfall happens.] If Ksh 250 is too expensive for you, you can decrease the units of insurance you buy. [If a farmer says he/she wants to buy 4 units of insurance] Ok, if you buy 4 units then now you will pay Ksh 200 but, your maximum payout when rainfall is bad is Ksh 20000. Are you comfortable to pay Ksh 200 to be insured at maximum Ksh 20000? [Same procedure goes on until the farmer decides.]

Now, the price of 1 unit of insurance is Ksh 100. You bought 4 units when the price was Ksh 50, because your maximum amount of insured was Ksh 20000. If you still want to buy 4 units then now you need to pay Ksh 400 because the price is now Ksh 100 per unit. Are you comfortable to pay Ksh 400 to be insured at maximum Ksh 20000? [In this moment, ASSISTANT should emphasize that Ksh 20000 is the maximum amount that the farmer can receive when bad rainfall happens.] If Ksh 400 is too expensive for you, you can decrease the units of insurance you buy. [If a farmer says he/she wants to buy 3 units of insurance] Ok, if you buy 3 units then now you will pay Ksh 300, but your maximum payout when rainfall is bad is Ksh 15000. Are you comfortable to pay Ksh 300 to be insured at maximum Ksh 15000?" [Same procedure goes on until the farmer decides not to buy any insurance.]

Now that you have made your decisions, we will determine the actual price of WII insurance for sorghum. One farmer will draw a price card for the sorghum insurance product. Remember that these prices are subsidized by the research team from Nairobi. Some groups will receive higher subsidies that others it all depends on what price we draw out of the hat. If you buy insurance next year, it will not be subsidized.

[GM: Have a farmer pull out the card from the envelope for Price of WII insurance for sorghum. Hold up the card and announce the price.]

The drawn price of 1 unit insurance for sorghum is [P]. If you said you would like to purchase at least one unit of the WII insurance for sorghum at this price, you purchase that number of units of the insurance by paying [P] per unit. If you said you did not want to buy any quantity of the WII insurance for sorghum at this price, you will not buy any WII insurance.

Now, for those who end up purchasing insurance products, your assistants will collect your contact information to notify Acre of your purchase decision. The agents of Acre will visit you later to help you complete the contract. We will also leave you with an informational pamphlet about index insurance with the price and quantity you agreed to buy so that you can remember the commitment you made.

We only have one auction left. This is for a product that does not yet exist for farmers in your area, but that Acre is considering developing. We therefore want to know how much you value this product.

3. Hypothetical Auction for AYI

Now you are going to bid hypothetically on a slightly Yield Insurance (AYI). Area Yield Insurance is similar to Weather Index Insurance, but is based on the actual yields of an area instead of rainfall outcomes. While Weather Index Insurance uses overall rainfall for your squared area to determine if you receive a payout, Area Yield Insurance uses the overall yield of farms in your squared area to determine whether you receive a payout.

If your areas overall yield for that season is bad, you receive a payout if you purchased insurance at the beginning of that season. Whether you receive a payout does not depend on the yield of your farm, but on the overall yield of your area. While Weather Index Insurance only protects you against drought, AYI protects you against any event that damages the yields of an area. This could be pests, disease, drought, hail, etc. This insurance product is not yet available because it takes time to design, so you will not make an actual purchase based on the results of this auction, but you should bid realistically because the results of this auction may help develop future AYI products.

[The subsequent process is same as the auction for WII except that the auction for AYI is hypothetical.]

4. Purchase and Closing

We thank you very much for your time and interest throughout the exercise. Please turn in your bid sheets to your assistant. Your assistants will ask some final questions [Second survey] and complete the fully binding contract to purchase what you agreed to in the auction. You are free to go when you complete the second survey and the contract.

[To the 10x10 participants only]: Recall that the insurance product you bid on today was based on square areas 10 km by 10 km. In reality, ACRE does not sell this product; we were using it as part of our experiment. You will actually be sold a product which is identical in every way, but the squares are 5 km by 5 km. This means that the area in which you are grouped to receive insurance payouts is smaller, and that you are more likely to receive a payout when you have a drought. Therefore, you will receive a slightly better product at the same price you agreed to in the auction.

III. Sorghum Insurance Policy: Chirps 5km x 5 km

Illustration:
Satellite Pixel/Location: Point 1

- **For how long will the cover last?** 25 pentads (approx. 125 days)

Germination Drought Cover

When does the cover start? *Pentad 61*
When does this cover end? *Pentad 64*
For how long will the cover run? *4 pentads*
When does a loss occur? *When the total rainfall recorded in any period of 5 days (pentad) is less than or equal to 14mm.*
How will the payable loss be calculated? *The loss payable per pentad is 6.25% of the total sum insured*
What is the maximum loss payable at this stage? *25% of the total sum insured.*

Vegetative Drought Cover

When does the cover start? *Pentad 65*
When does the cover end? *Pentad 70*
For how long will the cover run? *6 pentads*
When does a loss occur? *When the total rainfall recorded in any pentad is less than or equal to 4mm.*
How will the payable loss be calculated? *The loss payable per pentad is 5% of the total sum insured*
What is the maximum loss payable at this stage? *30% of the total sum insured*

Flowering Drought Cover

When does the cover start? *Pentad 71*
When does this cover end? *Pentad 6 of the following year*
For how long will the cover run? *8 pentads*
When does a loss occur? *When the total rainfall recorded in any pentad is less than or equal to 1mm.*
How will the payable loss be calculated? *The loss payable per pentad is 5% of the total sum insured*
What is the maximum loss payable at this stage? *40% of the total sum insured.*

Pre-harvest Excessive Rainfall Cover

When does the cover start? *Pentad 7 of the following year*
When does the cover end? *Pentad 13 of the following year*
For how long will the cover run? *7 pentads*
When does a loss occur? *When the total rainfall recorded in any pentad is greater than or equal to 7mm.*
How will the payable loss be calculated? *The loss payable per pentad is 3.57% of the total sum insured*
What is the maximum loss payable at this stage? *25% of the total sum insured*

Total Compensation & Deductible

The total loss is the sum of the 4 covers, with the maximum total loss allowed being 100% of the sum insured
An excess equal to 10% of the total sum insured applies.

The table below shows the triggers for all pixels in Tharaka Nithi County.

POINTID	Longitude	Latitude	Germination Trigger (mm)	Vegetative Trigger (mm)	Flowering Trigger (mm)	Pre-harvest Trigger (mm)
1	37.982785	0.024529	14.00	4.00	1.00	7.00
2	38.032785	0.024529	11.00	3.00	1.00	6.00
3	38.082785	0.024529	11.00	3.00	1.00	5.00
4	38.132785	0.024529	11.00	3.00	1.00	5.00
5	37.882785	-0.025471	16.00	3.00	1.00	10.00
6	37.932785	-0.025471	15.95	3.00	1.00	7.00
7	37.982785	-0.025471	15.00	3.00	1.00	5.90
8	38.032785	-0.025471	12.00	3.00	1.00	5.00
9	38.082785	-0.025471	11.00	2.45	1.00	6.00
10	38.132785	-0.025471	10.00	2.45	1.00	5.00
11	38.182785	-0.025471	9.00	2.45	1.00	5.00
12	38.232785	-0.025471	9.00	2.00	1.00	4.00
13	38.282785	-0.025471	8.00	2.00	1.00	3.90
14	37.882785	-0.075471	14.00	3.00	1.00	7.00
15	37.932785	-0.075471	14.00	3.00	1.00	7.00
16	37.982785	-0.075471	13.00	3.00	1.00	5.00
17	38.032785	-0.075471	11.00	3.00	1.00	6.00
18	38.082785	-0.075471	10.95	2.45	1.00	6.00
19	38.132785	-0.075471	10.00	2.45	1.00	8.00
20	38.182785	-0.075471	9.00	3.00	1.00	5.00
21	38.232785	-0.075471	9.00	2.00	1.00	3.90
22	38.282785	-0.075471	8.00	2.00	1.00	3.00
23	37.882785	-0.125471	14.00	3.00	1.00	7.00
24	37.932785	-0.125471	13.95	3.00	1.00	6.90
25	37.982785	-0.125471	13.00	3.00	1.00	5.00
26	38.032785	-0.125471	11.00	3.00	1.00	5.00
27	38.082785	-0.125471	10.95	2.45	1.00	5.90
28	38.132785	-0.125471	10.00	2.45	1.00	7.00
29	38.182785	-0.125471	9.00	3.00	1.00	5.00
31	37.432785	-0.175471	22.00	4.00	2.25	19.80
32	37.482785	-0.175471	23.95	4.00	2.00	16.00
33	37.782785	-0.175471	16.00	3.45	1.00	7.90
34	37.832785	-0.175471	15.00	3.00	1.00	6.00
35	37.882785	-0.175471	14.00	3.00	1.00	6.90

36	37.932785	-0.175471	13.00	3.00	1.00	7.00
37	37.982785	-0.175471	12.00	3.00	1.00	5.00
38	38.032785	-0.175471	9.95	3.00	1.00	5.00
39	38.082785	-0.175471	10.00	3.00	1.00	5.00
40	38.132785	-0.175471	9.00	3.00	1.00	5.00
41	38.182785	-0.175471	9.00	3.00	1.00	4.00
42	37.432785	-0.225471	21.00	4.00	2.00	17.80
43	37.482785	-0.225471	21.95	4.00	2.00	13.90
44	37.532785	-0.225471	22.00	3.45	2.00	13.90
45	37.582785	-0.225471	21.00	4.00	1.00	10.00
46	37.632785	-0.225471	19.95	4.00	2.00	10.00
47	37.682785	-0.225471	18.95	4.00	1.00	8.90
48	37.732785	-0.225471	18.95	3.00	1.00	7.00
49	37.782785	-0.225471	16.00	3.00	1.00	9.00
50	37.832785	-0.225471	15.00	3.00	1.00	8.00
51	37.882785	-0.225471	14.95	3.00	1.00	7.00
52	37.932785	-0.225471	13.00	3.00	1.00	7.00
53	37.982785	-0.225471	11.95	3.00	1.00	6.00
54	38.032785	-0.225471	10.00	3.00	1.00	5.00
55	38.082785	-0.225471	10.00	3.00	1.00	4.90
56	38.132785	-0.225471	9.00	3.00	1.00	4.90
57	37.482785	-0.275471	21.95	4.00	2.00	12.00
58	37.532785	-0.275471	21.00	3.00	2.00	12.00
59	37.582785	-0.275471	20.00	3.00	2.00	11.00
60	37.632785	-0.275471	19.00	3.00	1.00	9.90
61	37.682785	-0.275471	18.95	3.00	1.00	7.90
62	37.732785	-0.275471	18.00	3.00	1.00	8.00
63	37.782785	-0.275471	16.00	2.00	1.00	8.00
64	37.832785	-0.275471	15.00	2.00	1.00	7.00
65	37.882785	-0.275471	15.00	2.00	1.00	6.00
66	37.932785	-0.275471	13.00	2.00	1.00	10.00
67	37.982785	-0.275471	12.00	2.00	1.00	6.00
68	37.532785	-0.325471	20.00	3.00	1.00	10.90
69	37.582785	-0.325471	19.95	2.45	1.00	10.00
70	37.632785	-0.325471	18.95	3.00	1.00	9.00
71	37.682785	-0.325471	18.00	3.00	1.00	7.90
72	37.732785	-0.325471	18.00	3.00	1.00	8.90
73	37.782785	-0.325471	16.00	2.00	1.00	7.00
74	37.832785	-0.325471	15.00	2.00	1.00	6.00
75	37.882785	-0.325471	14.95	2.00	1.00	6.00
76	37.932785	-0.325471	13.00	2.00	1.00	7.00
77	37.632785	-0.375471	19.00	2.45	1.00	10.00
78	37.682785	-0.375471	18.00	2.45	1.00	8.00

79	37.732785	-0.375471	18.00	3.00	1.00	9.90
80	37.782785	-0.375471	15.00	2.00	1.00	8.00
81	37.832785	-0.375471	13.95	2.00	1.00	7.00
82	37.932785	-0.375471	12.00	2.00	1.00	5.00
83	37.682785	-0.425471	18.95	2.45	1.00	10.00
84	37.732785	-0.425471	19.00	3.00	1.00	10.00
85	37.782785	-0.425471	15.00	2.00	1.00	8.00

"Any kind of exposure not specified under this policy remains fully excluded
These terms are subject to approval by the insurer

We confirm we have read and understood the above terms and conditions and agree to this index cover:

Name of Insured: **Signature:**
Date:

Appendix B

Chapter 2 Appendix

Structural Approach for GNE Estimates

The simplest model begins with the following assumptions. First, each household selects only one UPC out of X UPCs available in the market. Second, the probability of purchasing each UPC is identical as $\frac{1}{X}$ across all UPCs in the market. Third, household purchases are independent to each other where the probability of a household purchasing a UPC is not affected by others' purchasing the same UPC in a market.[1] Under these assumptions, we specify the number of unique UPCs that H households, select as follows.

$$X(H) = X[1 - (1 - \frac{1}{X})^H] \tag{B.1}$$

By assuming a functional form that $X(H)$ follows, we can estimate the total number of unique UPCs available in the market using UPC counts and frequency data recorded by sample households.

Now, we relieve the unrealistic first two assumptions of the simple model by allowing for different products to have different purchase probabilities. Given that the different probability for each of the *population* UPCs is generally unknown, Mao et al. (2005) consider that these probabilities are identical within groups but different across groups. In their model, the number of these groups (referred to as 'support points' in Mao et al. (2005) and 'incidence groups' in Handbury and Weinstein (2015)) are determined by data through a likelihood-based procedure.

[1] In this setting, the probability that we observe one of the H households purchasing a particular barcode product equals to one minus the probability that none of the H households selects the same product, which is $1 - (1 - \frac{1}{X})^H$.

Accordingly, we define q_{km} as a probability of a barcode product in group k being purchased by a household in market m, such that $q_{jm} = q_{km}$ for all barcode products $j \in$ group k. Suppose there are K number of groups in each market with different values of q_{km}, where $k = \{1, ..., K\}$. Keeping the third assumption about the independence of household purchases, we can express the probability of any barcode products within group k purchased by one of H_m households in market m as $1-(1-q_{km})^{H_m}$. We take the weighted average of these probabilities across K groups using group shares α_{km} as weights to obtain the unconditional probability that a sample of H_m households purchases any one barcode available in market m. If market m has X_m unique barcode products available, similar to equation (B.1) the number of unique UPCs that a sample of H_m households purchases is as follows:

$$X_m(H_m) = X_m \sum_{k=1}^{K} \alpha_{km}(1-(1-q_{km})^{H_m}) \tag{B.2}$$

According to Mao et al. (2004), $X_m(H_m)$ should follow the negative exponential model where equation (B.2) is replaced by the following form.

$$X_m(H_m) = X_m \sum_{k=1}^{K} \alpha_{km}(1-\mathrm{e}^{\beta_{km} H_m}) \quad \text{where} \quad \beta_{km} = -\ln(1-q_{km}) \tag{B.3}$$

Equation (B.2) reduces to equation (B.1) when every UPC has the homogenous selection probability, or $q = \frac{1}{X}$.

Using a maximum likelihood methodology introduced in Mao et al. (2005), we can estimate α_{km}, q_{km}, and therefore X_m for any given K from $X_m(H_m)$ (i.e., the number of unique barcode products purchased by our sample in market m that we observe). We first obtain these parameters for a range of values for K as well as associated values of maximum likelihood functions for our 69 markets. We then calculate the Akaike Information Criterion (AIC) for each value of K in each market and choose one value of K that maximizes the sum of the AICs across all markets.

We find that 30 incidence groups are best suited for our purchase frequency data. Therefore, we estimate the total number of varieties X_m for each market using the parameters associated with $K = 30$. Given estimated parameters for α_{km} and q_{km}, we simply use equation (B.2) to recover X_m, which we refer to as "generalized negative exponential (GNE) estimates".

Appendix C

Chapter 3 Appendix

Derivation of Food Demand Elasticities

Since the EASI demand is a Hicksian demand system, we first obtain Hicksian price elasticity along with total expenditure elasticity and then recover Marshallian price elasticity. For ease of understanding the derivation, we display the two-way approximate EASI demand system below. Note that the household subsript is withheld for notational simplicity.

$$\begin{aligned} &w_i = \mu_i + \sum_{j=1}^{J} \alpha_{ij} \log p_j + \sum_{r=1}^{L} \beta_{ir} y^r + \sum_{j=1}^{J} \alpha_{ijy} (y \times \log p_j) + \sum_{k=1}^{K} \gamma_{ik} z_k + u_i, \\ &(i = 1, ...J-1) \quad \text{where} \quad y = \log x - \sum_{j=1}^{J} w_j \log p_j \end{aligned} \tag{C.1}$$

From the above equation, we first take partial derivatives with respect to $\log p_j$. This gives the following Hicksian semi elasticity.

$$\frac{\partial w_i}{\partial \log p_j} = \alpha_{ij} + \alpha_{ijy} y \tag{C.2}$$

Because $w_i = {q_i}^H p_i / x^H$ where the superscript H emphasizes variables are compensated, we can express the Hicksian semi elasticity as a function of the conventional Hicksian price elasticity.

$$\begin{aligned} \frac{\partial w_i}{\partial \log p_j} &= \frac{\partial ({q_i}^H p_i / x^H)}{\partial \log p_j} = \alpha_{ij} + \alpha_{ijy} y \\ &= \frac{\partial {q_i}^H}{\partial \log p_j} \frac{p_i}{x^H} + \frac{\partial p_i}{\partial \log p_j} \frac{{q_i}^H}{x^H} - \frac{{q_i}^H p_i}{(x^H)^2} \frac{\partial x^H}{\partial \log p_j} \\ &= \frac{\partial {q_i}^H}{\partial \log p_j} \frac{{q_i}^H}{{q_i}^H} \frac{p_i}{x^H} + \frac{\partial p_i}{\partial \log p_j} \frac{p_i}{p_i} \frac{{q_i}^H}{x^H} - \frac{{q_i}^H p_i}{(x^H)^2} \frac{\partial x^H}{\partial p_j} \frac{p_j}{1} \\ &= \frac{\partial \log {q_i}^H}{\partial \log p_j} w_i + 1_{ij} w_i - w_i w_j \end{aligned} \tag{C.3}$$

Dividing both sides of the last equation by w_i and rearranging the terms gives the Hicksian price elasiticity when $i = j$.

$$h_{ij} = \frac{\alpha_{ij} + \alpha_{ijy} y}{w_i} - 1_{ij} + w_j \tag{C.4}$$

The formula for the Hicksian price elasticity when $i \neq j$ is $h_{ij} = \frac{\alpha_{ij}+\alpha_{ijy}y}{w_i} + w_j$.

Next, we take partial derivatives of a matrix formation of equation (C.1) with respect to nominal total expenditure $\log x$, accounting for the budget share w_i appears on both sides of the demand equation. This gives the following $J \times 1$ (J equals to 20 in our study) vector of semi-expenditure elasticities, se.

$$se = (I_J + TP')^{-1} P \tag{C.5}$$

where I_J is the J dimension of the identity matrix, T is a $J \times 1$ vector whose ith element equals $\sum_{r=1}^{L} r\beta_{ir} y^{r-1}$, P is the $J \times 1$ vector of log prices. Using the relation between semi-expenditure elasticity and total expenditure elasticity (i.e., $e_i = \frac{se_i}{w_i} + 1$), we calculate the $J \times 1$ vector of total expenditure elasticities, e as

$$e = (diag(W))^{-1} \left[(I_J + TP')^{-1} T \right] + 1_J \tag{C.6}$$

where W is the $J \times 1$ vector of observed budget shares and 1_J is a $J \times 1$ vector of ones.

Once we have the Hicksian (compensated) price and total expenditure elasticities, the Marshallian (uncompensated) price elasticity can be easily calculated from the Slutsky equation.

$$e_{ij} = h_{ij} - w_j e_i \tag{C.7}$$

where e_i is a total expenditure elasticity for demand category i, or ith element of the total expenditure vector e.

For zero demands (i.e., censoring), we calculate predicted budget shares (i.e., conditional means of observed budget shares) and replace the observed budget shares with them in the above equations to obtain expected demand elasticities.

Table C.1: Price and Total Expenditure Elasticities for the First Quartile: Panel Model

	Food Category																			Expenditure
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. Rice	-3.55*	0.70*	-0.17	0.43	-0.40	-0.13	0.60*	-0.64*	0.23*	-0.22	0.49*	-0.78*	-0.10	-0.75*	0.60*	-0.21	0.07	-0.08	-0.05*	2.33*
	(-5.02)	(2.51)	(-1.46)	(1.56)	(-1.55)	(-1.49)	(2.74)	(-3.00)	(2.33)	(-1.75)	(2.13)	(-2.68)	(-0.94)	(-3.23)	(2.75)	(-1.90)	(1.91)	(-0.89)	(-2.35)	(4.41)
2. Maize	0.40*	-1.09*	0.41*	-0.10	0.07	0.18*	-0.05	-0.36	-0.05	0.33	-0.08	0.36	0.02	0.12	0.06	0.04	0.06*	0.02	0.02*	0.30
	(2.06)	(-4.01)	(2.25)	(-0.40)	(0.47)	(2.61)	(-1.03)	(-1.87)	(-1.08)	(1.83)	(-0.88)	(1.57)	(0.60)	(1.48)	(1.34)	(1.31)	(2.42)	(0.51)	(2.22)	(0.23)
3. Wheat†	-0.24	1.15*	-4.05*	0.06	1.30*	-0.06	-0.13	0.41	0.07	-0.04	-0.76*	-0.53*	-0.02	0.02	0.03	0.03	0.09*	-0.49*	-0.07*	2.06*
	(-1.50)	(3.42)	(-6.27)	(0.03)	(3.74)	(-0.69)	(-0.76)	(1.70)	(0.59)	(-0.38)	(-3.20)	(-2.34)	(-0.17)	(-0.10)	(0.22)	(0.30)	(2.43)	(-3.06)	(-2.94)	(4.80)
4. Cassava	0.44*	-0.24	0.08	-1.39*	-0.14	0.00	-0.27*	-0.20	-0.06	-0.13	0.70*	0.10	0.10	0.17	0.23	0.09	0.01	-0.25*	0.02	0.64
	(2.08)	(-0.90)	(0.96)	(-4.61)	(-0.59)	(-0.12)	(-2.16)	(-1.07)	(-0.88)	(-0.57)	(2.32)	(0.22)	(1.23)	(1.45)	(1.63)	(1.19)	(0.42)	(-2.06)	(1.67)	(1.95)
5. Roots†	-0.32	0.20	0.95*	-0.17	-2.74*	-0.09	-1.19*	0.56	0.42*	0.11	0.16	0.76	-0.16	0.01	0.51*	-0.02	-0.10*	-0.07	-0.03	0.50
	(-1.07)	(0.36)	(2.84)	(-0.43)	(-4.18)	(-1.43)	(-2.84)	(1.91)	(2.63)	(0.87)	(0.90)	(1.52)	(-1.23)	(0.30)	(2.48)	(-0.08)	(-2.42)	(-0.62)	(-1.55)	(0.94)
6. Sugar	-0.23	1.02*	-0.07	-0.03	-0.25	-0.89*	0.54*	-0.05	0.14	-0.14	-0.00	-0.19	-0.30*	0.03	0.30*	0.22*	-0.05	0.26*	-0.03	1.05*
	(-1.43)	(4.77)	(-0.33)	(-0.32)	(-1.71)	(-5.00)	(3.01)	(-0.38)	(1.62)	(-1.48)	(0.24)	(-1.09)	(-2.46)	(0.24)	(2.94)	(2.26)	(-1.15)	(2.53)	(-1.16)	(4.94)
7. Pulses	0.72*	-0.17	-0.06	-0.36*	-1.35*	0.25*	-1.71*	0.04	0.19*	-0.11	0.40*	0.09	0.48*	0.63*	0.14	-0.02	0.14*	-0.22*	0.02	0.48
	(3.16)	(-1.29)	(-0.15)	(-2.29)	(-3.70)	(2.93)	(-6.20)	(0.63)	(2.37)	(-0.67)	(2.71)	(0.80)	(3.45)	(3.54)	(1.33)	(-0.27)	(2.96)	(-2.05)	(1.12)	(1.87)
8. Nuts†	-1.00*	-1.95*	0.46	-0.49	0.82	-0.06	-0.05	-2.41	-0.08	0.36	-0.40	-1.94*	-0.43*	-0.04	-0.21	0.08	-0.00	0.04	0.08*	2.55*
	(-2.27)	(-2.14)	(1.44)	(-1.48)	(1.61)	(-0.58)	(-0.37)	(-3.88)	(-0.76)	(0.97)	(-1.25)	(-2.29)	(-1.98)	(-0.60)	(-1.14)	(0.90)	(0.01)	(0.16)	(2.33)	(2.91)
9. Vegetables	0.34*	-0.20*	0.12*	-0.09	0.38*	0.06	0.14*	0.03	-0.60	-0.20*	0.02	-0.03	-0.14*	0.27*	-0.18*	0.01	0.11*	0.19*	-0.04*	0.72*
	(4.57)	(-2.42)	(2.03)	(-0.81)	(3.93)	(1.83)	(2.43)	(0.44)	(-9.60)	(-2.99)	(0.10)	(-0.11)	(-3.40)	(5.69)	(-4.17)	(0.06)	(5.73)	(3.70)	(-3.64)	(5.17)
10. Fruit	-0.36	1.12	-0.05	-0.35	0.10	-0.12	-0.27	0.39	-0.47*	-1.58*	0.45	-0.74	-0.24	-0.02	0.33	-0.03	-0.03	0.12	-0.01	2.21*
	(-1.25)	(1.69)	(-0.33)	(-0.90)	(0.06)	(-1.33)	(-1.10)	(1.01)	(-2.14)	(-4.15)	(1.38)	(-1.45)	(-1.50)	(-0.22)	(1.83)	(-0.26)	(-0.86)	(0.97)	(-0.66)	(2.64)
11. Red meat	0.50*	-0.65	-0.67*	0.83*	0.10	-0.04	0.31*	-0.30	-0.14	0.28	-1.67*	-0.90*	0.30*	-0.10	-1.03*	-0.32*	0.05	-0.34*	-0.06*	2.81*
	(1.98)	(-1.83)	(-3.92)	(2.57)	(0.45)	(-0.37)	(2.02)	(-1.81)	(-1.36)	(0.90)	(-4.93)	(-2.66)	(2.37)	(-1.04)	(-4.55)	(-4.13)	(1.86)	(-2.49)	(-2.83)	(5.72)
12. Poultry	-0.80*	0.07	-0.45*	-0.04	0.53	-0.14	-0.11	-1.17*	-0.30	-0.47*	-0.78*	-3.10*	-0.56*	-0.04	-0.39*	0.12	-0.04	0.40*	-0.00	4.32*
	(-3.02)	(0.10)	(-2.47)	(-0.48)	(0.95)	(-1.47)	(-0.63)	(-3.14)	(-1.77)	(-2.16)	(-2.25)	(-4.02)	(-3.10)	(-0.63)	(-2.05)	(1.62)	(-1.07)	(2.04)	(0.08)	(3.59)
13. Eggs	-0.62	-0.30	-0.11	0.48	-0.94	-0.69*	2.17*	-1.40*	-1.05*	-0.79*	1.31*	-3.13*	-0.94	0.41	0.44	-0.86*	0.06	0.25	0.02	3.75*
	(-1.24)	(-0.33)	(-0.28)	(0.68)	(-1.73)	(-2.36)	(3.71)	(-2.83)	(-3.12)	(-2.53)	(2.20)	(-3.87)	(-1.94)	(1.24)	(1.09)	(-2.55)	(0.44)	(0.79)	(0.66)	(4.06)
14. Fish†	-0.94*	0.25	0.05	0.24	-0.03	0.01	0.75*	0.00	0.34*	0.01	-0.05	0.10	0.14	-1.52*	-0.04	0.14	-0.01	-0.28*	-0.08*	1.37*
	(-2.94)	(1.54)	(0.36)	(1.24)	(-0.48)	(0.08)	(3.12)	(-0.22)	(3.05)	(-0.07)	(-0.30)	(0.22)	(1.64)	(-8.29)	(-0.28)	(1.57)	(-0.58)	(-2.38)	(-3.04)	(4.38)
15. Dairy	0.95*	0.07	0.02	0.39	0.82*	0.18*	0.15	-0.21	-0.41*	0.32*	-1.44*	-0.64*	0.15	-0.07	-0.91*	-0.19*	0.03	0.13	-0.08*	1.75*
	(3.93)	(0.68)	(-0.21)	(1.48)	(3.46)	(3.06)	(1.01)	(-1.63)	(-4.68)	(2.19)	(-5.39)	(-2.57)	(1.13)	(-0.68)	(-3.46)	(-2.86)	(1.32)	(1.01)	(-4.39)	(5.77)
16. Fats†	-0.38	0.18	0.11	0.24	-0.09	0.22	-0.06	0.20	0.01	0.01	-0.54*	0.60	-0.37*	0.27	-0.24	-0.76	0.25*	0.09	0.17*	0.73
	(-1.42)	(1.10)	(0.86)	(1.11)	(-0.33)	(1.86)	(-0.36)	(1.18)	(0.02)	(0.17)	(-2.22)	(1.94)	(-2.02)	(1.53)	(-1.85)	(-1.63)	(2.35)	(0.88)	(2.50)	(1.52)
17. Coffee†	0.62*	1.32*	0.59*	0.06	-0.93*	-0.18	1.11*	0.01	1.03*	-0.16	0.40*	-0.25	0.12	-0.07	0.19	0.88*	-0.74*	0.82*	0.11*	1.09*
	(2.20)	(4.35)	(2.70)	(0.24)	(-3.20)	(-1.14)	(3.19)	(0.09)	(4.24)	(-0.91)	(2.04)	(-0.75)	(0.58)	(-0.48)	(1.35)	(3.83)	(-8.09)	(3.90)	(2.07)	(3.24)
18. Soft drink†	-0.26	0.18	-1.14*	-1.00*	-0.27	0.34*	-0.68*	0.08	0.68*	0.24	-0.96*	1.48*	0.17	-0.67*	0.26	0.12	0.31*	-1.13*	0.01	1.17
	(-0.81)	(0.35)	(-2.71)	(-2.07)	(-0.69)	(2.27)	(-2.11)	(0.36)	(2.77)	(1.09)	(-2.00)	(2.16)	(0.82)	(-2.35)	(1.04)	(0.94)	(3.01)	(-3.18)	(0.21)	(1.65)
19. Other food	-0.26*	0.35*	-0.33*	0.22*	-0.26*	-0.08	0.11	0.48	-0.29*	0.03	-0.25*	0.24*	0.08	-0.38*	-0.30*	0.52*	0.10*	0.02	-0.88*	0.56*
	(-1.96)	(3.39)	(-3.01)	(2.06)	(-1.98)	(-1.04)	(1.16)	(5.03)	(-3.22)	(0.73)	(-2.01)	(2.30)	(1.36)	(-3.75)	(-3.72)	(4.39)	(2.20)	(0.24)	(-9.99)	(2.62)

Notes: This table shows consumption quantity elasticities with respect to food prices and total expenditure with t-value at median in parenthesis (* $p < 0.05$). We estimate these elasticities for the first quartile sample with the panel model. Refer to the following full name of each food category for row title with †; 1. Rice 2. Maize 3. Wheat and other cereals 4. Cassava 5. Roots, Tubers and, other starches 6. Sugar 7. Pulses 8. Nuts and seeds 9. Vegetables 10. Fruit 11. Red meat 12. Poultry 13. Eggs 14. Fish and seafood 15. Dairy 16. Fats and oils 17. Coffee, tea, and cocoa 18. Soft drink and juice 19. Other food Expenditure

Table C.2: Price and Total Expenditure Elasticities for the Second Quartile: Panel Model

	Food Category																			Expenditure
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. Rice	-2.49*	0.61*	-0.13	0.43*	-0.31	-0.04	0.39*	-0.31*	0.15*	-0.01	0.35*	-0.36	-0.00	-0.42*	0.29*	-0.18*	0.05	-0.05	-0.03*	1.36*
	(-5.43)	(2.38)	(-1.38)	(1.97)	(-1.67)	(-0.85)	(2.56)	(-2.11)	(2.02)	(-0.29)	(2.25)	(-1.88)	(0.07)	(-2.72)	(2.18)	(-2.26)	(1.90)	(-0.81)	(-2.20)	(4.81)
2. Maize	0.35	-1.08*	0.35*	-0.19	0.09	0.17*	-0.05	-0.44*	-0.07	0.25	-0.03	0.11	0.02	0.09	0.05	0.07	0.06*	0.07	0.02*	0.71
	(1.84)	(-3.63)	(2.11)	(-1.14)	(0.61)	(2.50)	(-0.83)	(-2.02)	(-1.54)	(1.46)	(-0.32)	(0.80)	(0.64)	(1.13)	(1.31)	(1.72)	(2.35)	(1.28)	(2.24)	(1.48)
3. Wheat†	-0.19	0.92*	-2.92*	0.32	0.93*	-0.11	-0.02	0.51*	0.05	-0.02	-0.62*	-0.06	0.00	0.15	-0.11	-0.01	0.07*	-0.38*	-0.04*	1.39*
	(-1.47)	(3.05)	(-6.39)	(1.79)	(3.36)	(-1.65)	(-0.07)	(2.54)	(0.38)	(-0.19)	(-3.21)	(-0.56)	(0.02)	(1.54)	(-1.50)	(-0.25)	(2.34)	(-3.00)	(-2.29)	(5.16)
4. Cassava	0.52*	-0.32	0.27*	-1.46*	-0.28	-0.01	-0.27*	-0.10	0.01	-0.02	0.70*	0.09	0.10	0.18	0.32*	0.10	-0.02	-0.22*	0.03*	0.17
	(2.33)	(-1.15)	(2.11)	(-4.69)	(-1.23)	(-0.28)	(-2.31)	(-0.35)	(0.14)	(0.13)	(2.44)	(0.24)	(1.48)	(1.54)	(2.11)	(1.38)	(-1.02)	(-2.10)	(2.03)	(1.08)
5. Roots†	-0.32	0.12	0.68*	-0.34	-2.43*	-0.08	-0.85*	0.31	0.28	-0.10	0.20	0.32	-0.16	-0.17	0.41*	0.05	-0.07*	-0.04	-0.02	1.26*
	(-1.47)	(0.41)	(2.69)	(-1.42)	(-4.70)	(-1.57)	(-2.92)	(1.49)	(2.50)	(-0.58)	(1.35)	(0.91)	(-1.69)	(-1.36)	(2.45)	(1.00)	(-2.17)	(-0.47)	(-1.21)	(2.87)
6. Sugar	-0.09	0.79*	-0.18	-0.08	-0.18	-0.74*	0.39*	-0.02	0.13	-0.11	-0.03	-0.19	-0.16*	-0.00	0.16*	0.15*	-0.06	0.16*	-0.02	1.10*
	(-0.75)	(4.59)	(-1.81)	(-1.05)	(-1.71)	(-5.07)	(2.81)	(-0.24)	(1.89)	(-1.41)	(-0.20)	(-1.57)	(-2.01)	(-0.15)	(2.43)	(2.12)	(-1.86)	(2.31)	(-1.12)	(6.74)
7. Pulses	0.56*	-0.13	0.00	-0.40*	-1.06*	0.22*	-1.58*	-0.04	0.16*	-0.07	0.28*	-0.03	0.37*	0.55*	0.18	-0.07	0.11*	-0.17	-0.02	0.75
	(2.83)	(-0.96)	(0.29)	(-2.71)	(-3.56)	(2.65)	(-6.17)	(-0.23)	(2.19)	(-0.39)	(2.25)	(0.05)	(3.08)	(3.37)	(1.93)	(-1.17)	(2.60)	(-1.91)	(-1.11)	(0.28)
8. Nuts†	-0.61	-1.84*	0.66*	-0.22	0.57	-0.02	-0.09	-1.99*	-0.02	0.45	-0.50	-1.21	-0.39*	0.21	-0.32	0.07	0.00	-0.05	0.09*	1.39*
	(-1.70)	(-2.32)	(2.01)	(-0.85)	(1.38)	(-0.31)	(-0.64)	(-3.93)	(-0.32)	(1.48)	(-1.63)	(-1.78)	(-2.02)	(1.17)	(-1.62)	(0.78)	(0.10)	(-0.27)	(2.46)	(2.78)
9. Vegetables	0.24*	-0.17*	0.08	-0.03	0.38*	0.08*	0.15*	0.01	-0.64*	-0.22*	0.08	0.02	-0.08*	0.23*	-0.12*	-0.03	0.08*	0.16*	-0.04*	0.64*
	(3.74)	(-2.31)	(1.52)	(-0.09)	(4.17)	(2.27)	(2.58)	(0.36)	(-11.26)	(-3.54)	(1.16)	(0.64)	(-2.42)	(5.29)	(-3.21)	(-0.99)	(4.87)	(3.63)	(-3.76)	(5.63)
10. Fruit	0.00	0.83	-0.01	-0.08	-0.14	-0.07	-0.09	0.43	-0.32*	-1.10*	0.50	-0.31	-0.05	0.19	0.23	-0.02	-0.02	0.11	-0.02	0.94*
	(-0.05)	(1.51)	(-0.08)	(-0.29)	(-0.55)	(-1.06)	(-0.67)	(1.42)	(-2.16)	(-4.07)	(1.68)	(-0.86)	(-0.59)	(1.07)	(1.63)	(-0.29)	(-0.62)	(1.01)	(-1.13)	(2.41)
11. Red meat	0.33	-0.29	-0.50*	0.65*	0.17	-0.04	0.16	-0.30*	-0.02	0.28	-1.56*	-0.58*	0.27*	-0.11	-0.85*	-0.22*	0.07*	-0.31*	-0.05*	2.17*
	(1.86)	(-1.36)	(-4.09)	(3.08)	(1.09)	(-0.82)	(1.53)	(-2.39)	(-0.35)	(1.57)	(-7.07)	(-2.61)	(3.52)	(-1.36)	(-5.21)	(-4.09)	(3.39)	(-3.34)	(-3.40)	(7.36)
12. Poultry	-0.41*	-0.04	-0.07	0.02	0.23	-0.10	-0.09	-0.60*	-0.07	-0.21	-0.51*	-1.94*	-0.33*	0.11	-0.21	0.10	-0.04	0.22	0.02	2.27*
	(-2.24)	(-0.31)	(-0.82)	(-0.07)	(0.51)	(-1.58)	(-0.74)	(-2.44)	(-0.96)	(-1.43)	(-2.11)	(-4.02)	(-2.65)	(0.73)	(-1.56)	(1.47)	(-1.48)	(1.72)	(1.55)	(3.73)
13. Eggs	-0.09	-0.02	-0.03	0.43	-0.87*	-0.36	1.34*	-1.05*	-0.45*	-0.20	1.25*	-1.83*	-0.83*	0.35	0.11	-0.55*	0.08	0.15	-0.02	2.30*
	(-0.28)	(0.23)	(-0.11)	(1.16)	(-2.05)	(-1.90)	(3.17)	(-2.88)	(-2.40)	(-1.10)	(2.94)	(-3.41)	(-2.81)	(1.47)	(0.37)	(-2.41)	(1.05)	(0.68)	(-0.48)	(4.37)
14. Fish†	-0.61*	0.24	0.17	0.22	-0.21	0.00	0.61*	0.17	0.28*	0.16	-0.08	0.25	0.12	-1.36*	0.04	0.02	-0.02	-0.19*	-0.05*	0.90*
	(-2.50)	(1.35)	(1.77)	(1.33)	(-1.41)	(-0.03)	(2.97)	(1.49)	(2.94)	(1.32)	(-0.66)	(1.41)	(1.75)	(-9.45)	(0.75)	(0.31)	(-0.96)	(-2.12)	(-2.82)	(4.36)
15. Dairy	0.46*	0.07	-0.13*	0.44*	0.64*	0.09*	0.17	-0.28*	-0.21*	0.20	-1.25*	-0.35*	0.04	0.02	-1.07*	-0.12*	0.02	0.02	-0.04*	1.55*
	(2.61)	(0.82)	(-2.08)	(2.65)	(3.73)	(2.28)	(1.84)	(-2.60)	(-3.98)	(1.62)	(-6.58)	(-2.10)	(0.28)	(0.14)	(-5.92)	(-2.69)	(1.20)	(-0.04)	(-3.77)	(7.51)
16. Fats†	-0.47	0.32	0.00	0.20	0.15	0.16	-0.15	0.12	-0.09	-0.03	-0.46*	0.39	-0.27	0.02	-0.17	-0.57	0.28*	-0.00	0.09*	0.99*
	(-1.80)	(1.61)	(-0.00)	(1.15)	(1.08)	(1.66)	(-1.11)	(0.92)	(-1.19)	(-0.23)	(-2.21)	(1.59)	(-1.88)	(0.18)	(-1.69)	(-0.84)	(2.49)	(0.05)	(2.03)	(2.53)
17. Coffee†	0.40*	0.98*	0.42*	-0.19	-0.53*	-0.21	0.67*	0.01	0.49*	-0.11	0.60*	-0.31	0.14	-0.12	0.11	0.83*	-0.94*	0.43*	0.08*	1.24*
	(2.10)	(4.14)	(2.59)	(-1.52)	(-2.81)	(-1.95)	(2.79)	(0.09)	(3.48)	(-0.76)	(3.23)	(-1.51)	(1.14)	(-1.16)	(1.29)	(4.48)	(-13.89)	(3.31)	(2.06)	(4.98)
18. Soft drink†	-0.23	0.30	-0.92*	-0.78*	-0.17	0.19	-0.49*	-0.09	0.38*	0.18	-0.96*	0.81	0.10	-0.48*	0.03	-0.03	0.17*	-1.10*	0.02	1.86*
	(-0.87)	(0.86)	(-2.72)	(-2.23)	(-0.62)	(1.92)	(-2.06)	(-0.47)	(2.40)	(0.85)	(-2.68)	(1.72)	(0.68)	(-2.22)	(0.04)	(-0.22)	(2.58)	(-4.16)	(1.08)	(3.53)
19. Other food	-0.30*	0.46*	-0.28*	0.25*	-0.13	-0.08	-0.14	0.52*	-0.32*	-0.10	-0.35*	0.30*	-0.03	-0.37*	-0.22*	0.37*	0.10*	0.10	-0.91*	0.57*
	(-2.20)	(4.11)	(-2.35)	(2.55)	(-1.01)	(-0.99)	(-1.09)	(5.16)	(-3.43)	(-0.87)	(-2.82)	(2.69)	(-0.13)	(-3.85)	(-3.10)	(3.73)	(2.23)	(1.54)	(-12.10)	(2.94)

Notes: This table shows consumption quantity elasticities with respect to food prices and total expenditure with t-value at median in parenthesis (* $p < 0.05$). We estimate these elasticities for the second quartile sample with the panel model. Refer to the following full name of each food category for row title with †; 1. Rice 2. Maize 3. Wheat and other cereals 4. Cassava 5. Roots, Tubers and, other starches 6. Sugar 7. Pulses 8. Nuts and seeds 9. Vegetables 10. Fruit 11. Red meat 12. Poultry 13. Eggs 14. Fish and seafood 15. Dairy 16. Fats and oils 17. Coffee, tea, and cocoa 18. Soft drink and juice 19. Other food Expenditure

Table C.3: Price and Total Expenditure Elasticities for the Third Quartile: Panel Model

	Food Category																			Expenditure
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. Rice	-1.98*	0.51*	-0.11	0.42*	-0.28	0.00	0.26	-0.15	0.08	0.10	0.29*	-0.16	0.05	-0.27*	0.13	-0.18*	0.04	-0.04	-0.03	1.01*
	(-5.43)	(2.00)	(-1.12)	(1.97)	(-1.61)	(0.25)	(2.11)	(-1.15)	(1.38)	(0.81)	(2.04)	(-0.93)	(1.16)	(-2.11)	(1.43)	(-2.35)	(1.71)	(-0.64)	(-1.93)	(4.43)
2. Maize	0.32	-1.05*	0.31	-0.28	0.11	0.17*	-0.03	-0.54*	-0.08	0.18	0.01	-0.11	0.02	0.06	0.05	0.11	0.07*	0.11	0.02*	0.86*
	(1.52)	(-2.98)	(1.78)	(-1.33)	(0.70)	(2.39)	(-0.38)	(-1.99)	(-1.55)	(1.05)	(0.25)	(-0.37)	(0.62)	(0.79)	(1.14)	(1.89)	(2.24)	(1.55)	(2.22)	(1.96)
3. Wheat†	-0.16	0.71*	-2.19*	0.44*	0.70*	-0.14*	0.05	0.54*	0.02	-0.00	-0.53*	0.24	0.01	0.22*	-0.18*	-0.03	0.06*	-0.32*	-0.02	1.13*
	(-1.23)	(2.44)	(-6.04)	(2.20)	(2.75)	(-2.01)	(0.50)	(2.69)	(-0.01)	(-0.02)	(-2.86)	(1.08)	(0.17)	(2.03)	(-2.07)	(-0.64)	(2.00)	(-2.60)	(-1.51)	(4.94)
4. Cassava	0.60*	-0.44	0.44*	-1.53*	-0.42	-0.02	-0.29*	-0.02	0.06	0.06	0.71*	0.08	0.10	0.19	0.40*	0.10	-0.04	-0.21	0.03*	-0.01
	(2.28)	(-1.28)	(2.32)	(-4.29)	(-1.54)	(-0.44)	(-2.21)	(0.27)	(0.93)	(0.57)	(2.32)	(0.21)	(1.50)	(1.45)	(2.19)	(1.37)	(-1.69)	(-1.85)	(1.99)	(0.35)
5. Roots†	-0.34	0.12	0.51*	-0.43	-2.26*	-0.07	-0.61*	0.17	0.23*	-0.23	0.23	0.05	-0.17	-0.27	0.37*	0.10	-0.05	-0.02	-0.00	1.51*
	(-1.53)	(0.48)	(2.23)	(-1.72)	(-4.56)	(-1.46)	(-2.54)	(0.76)	(2.27)	(-1.27)	(1.45)	(-0.11)	(-1.75)	(-1.79)	(2.21)	(1.51)	(-1.74)	(-0.28)	(-0.29)	(3.42)
6. Sugar	-0.00	0.69*	-0.26*	-0.11	-0.15	-0.64*	0.32*	-0.01	0.13	-0.09	-0.06	-0.20	-0.08	-0.02	0.08	0.12	-0.07*	0.11	-0.02	1.08*
	(0.07)	(4.24)	(-2.37)	(-1.34)	(-1.49)	(-4.23)	(2.40)	(-0.04)	(1.90)	(-1.10)	(-0.53)	(-1.62)	(-1.11)	(-0.41)	(1.51)	(1.69)	(-1.98)	(1.91)	(-0.96)	(7.40)
7. Pulses	0.41*	-0.07	0.06	-0.42*	-0.82*	0.19*	-1.48*	-0.11	0.15	-0.04	0.16	-0.12	0.26*	0.49*	0.21*	-0.11	0.09*	-0.14	-0.04*	0.88*
	(2.31)	(-0.38)	(0.59)	(-2.58)	(-3.05)	(2.23)	(-5.59)	(-0.85)	(1.91)	(-0.16)	(1.32)	(-0.65)	(2.44)	(2.98)	(2.01)	(-1.50)	(2.08)	(-1.58)	(-2.04)	(3.84)
8. Nuts†	-0.30	-1.87*	0.79*	-0.04	0.34	0.00	-0.14	-1.67*	-0.01	0.52	-0.56	-0.63	-0.36	0.39	-0.39	0.05	0.00	-0.11	0.08*	0.95*
	(-0.99)	(-2.30)	(2.18)	(-0.13)	(0.98)	(0.04)	(-0.89)	(-3.71)	(-0.24)	(1.62)	(-1.79)	(-1.18)	(-1.87)	(1.72)	(-1.79)	(0.58)	(0.13)	(-0.68)	(2.29)	(2.34)
9. Vegetables	0.14*	-0.15	0.05	0.01	0.36*	0.09*	0.15*	0.00	-0.68*	-0.23*	0.13*	0.07	-0.03	0.20*	-0.06	-0.06	0.04*	0.13*	-0.04*	0.60*
	(2.50)	(-1.95)	(0.82)	(0.58)	(3.97)	(2.27)	(2.37)	(0.24)	(-11.21)	(-3.35)	(2.06)	(1.21)	(-0.69)	(4.40)	(-1.42)	(-1.71)	(3.05)	(3.17)	(-3.33)	(5.26)
10. Fruit	0.21	0.56	0.02	0.04	-0.30	-0.04	-0.02	0.44	-0.27*	-0.83*	0.53	-0.08	0.06	0.30	0.17	-0.02	-0.01	0.10	-0.03	0.56
	(0.90)	(1.19)	(0.08)	(0.21)	(-0.97)	(-0.72)	(-0.15)	(1.52)	(-2.03)	(-3.02)	(1.74)	(-0.16)	(0.66)	(1.43)	(1.32)	(-0.31)	(-0.29)	(0.94)	(-1.44)	(1.75)
11. Red meat	0.24	-0.10	-0.42*	0.53*	0.20	-0.04	0.07	-0.30*	0.03	0.28	-1.50*	-0.40*	0.25*	-0.11	-0.75*	-0.17*	0.09*	-0.29*	-0.04*	1.87*
	(1.47)	(-0.72)	(-3.65)	(3.00)	(1.44)	(-0.97)	(0.66)	(-2.59)	(0.61)	(1.93)	(-7.80)	(-2.34)	(3.66)	(-1.41)	(-4.98)	(-3.62)	(3.72)	(-3.42)	(-3.32)	(8.05)
12. Poultry	-0.18	-0.22	0.15	0.03	0.03	-0.08	-0.10	-0.30	0.01	-0.07	-0.36	-1.28*	-0.20*	0.20	-0.11	0.07	-0.04	0.12	0.02	1.56*
	(-1.19)	(-0.90)	(0.88)	(-0.03)	(-0.09)	(-1.49)	(-0.96)	(-1.59)	(0.10)	(-0.54)	(-1.74)	(-3.48)	(-2.13)	(1.50)	(-0.92)	(1.05)	(-1.61)	(1.22)	(1.83)	(3.50)
13. Eggs	0.21	0.08	0.02	0.38	-0.81*	-0.17	0.82*	-0.85*	-0.14	0.14	1.19*	-1.06*	-0.77*	0.30	-0.07	-0.36*	0.09	0.09	-0.05	1.76*
	(0.90)	(0.44)	(0.06)	(1.22)	(-2.03)	(-1.15)	(2.36)	(-2.67)	(-1.24)	(0.33)	(3.04)	(-2.69)	(-2.85)	(1.36)	(-0.46)	(-2.00)	(1.28)	(0.48)	(-1.48)	(4.29)
14. Fish†	-0.40	0.19	0.26*	0.20	-0.34	-0.00	0.51*	0.29*	0.22*	0.27	-0.10	0.36	0.11	-1.27*	0.09	-0.07	-0.02	-0.14	-0.04*	0.74*
	(-1.91)	(1.06)	(2.12)	(1.23)	(-1.90)	(-0.13)	(2.64)	(2.11)	(2.58)	(1.86)	(-0.84)	(1.89)	(1.64)	(-9.12)	(1.32)	(-1.18)	(-1.12)	(-1.73)	(-2.42)	(3.63)
15. Dairy	0.19	0.07	-0.22*	0.47*	0.55*	0.04	0.19*	-0.32*	-0.10*	0.13	-1.16*	-0.19	-0.02	0.07	-1.15*	-0.09*	0.01	-0.05	-0.02	1.44*
	(0.86)	(0.74)	(-2.92)	(2.98)	(3.38)	(1.13)	(2.11)	(-2.95)	(-2.19)	(0.97)	(-6.57)	(-1.32)	(-0.61)	(0.90)	(-6.35)	(-2.07)	(0.86)	(-0.94)	(-1.90)	(8.35)
16. Fats†	-0.53	0.47	-0.07	0.18	0.33	0.13	-0.22	0.07	-0.16	-0.06	-0.41*	0.26	-0.21	-0.15	-0.14	-0.42	0.30*	-0.07	0.04	1.09*
	(-1.95)	(1.79)	(-0.55)	(1.09)	(1.63)	(1.34)	(-1.33)	(0.55)	(-1.63)	(-0.45)	(-2.02)	(1.18)	(-1.61)	(-1.19)	(-1.36)	(-0.41)	(2.46)	(-0.92)	(1.23)	(2.73)
17. Coffee†	0.29	0.83*	0.33*	-0.33*	-0.33*	-0.23*	0.44*	0.01	0.21*	-0.08	0.71*	-0.33	0.15	-0.14	0.07	0.82*	-1.05*	0.23*	0.06	1.26*
	(1.72)	(3.71)	(2.18)	(-2.31)	(-2.11)	(-2.06)	(2.23)	(0.09)	(2.26)	(-0.54)	(3.48)	(-1.82)	(1.36)	(-1.39)	(1.00)	(4.26)	(-14.20)	(2.29)	(1.71)	(5.62)
18. Soft drink†	-0.20	0.43	-0.78*	-0.62*	-0.10	0.11	-0.35	-0.19	0.23	0.15	-0.96*	0.43	0.06	-0.35	-0.11	-0.11	0.09	-1.08*	0.04	1.95*
	(-0.83)	(1.22)	(-2.46)	(-2.11)	(-0.46)	(1.40)	(-1.78)	(-1.00)	(1.91)	(0.60)	(-2.72)	(0.82)	(0.37)	(-1.92)	(-0.90)	(-1.39)	(1.81)	(-4.27)	(1.61)	(4.00)
19. Other food	-0.32*	0.55*	-0.20	0.28*	0.01	-0.08	-0.35*	0.52*	-0.32*	-0.22	-0.42*	0.34*	-0.14	-0.33*	-0.11	0.18	0.10	0.17*	-0.95*	0.63*
	(-2.01)	(4.07)	(-1.33)	(2.47)	(0.39)	(-0.85)	(-2.36)	(4.58)	(-3.15)	(-1.95)	(-2.87)	(2.62)	(-1.43)	(-3.25)	(-1.49)	(1.72)	(1.83)	(2.18)	(-11.71)	(3.07)

Notes: This table shows consumption quantity elasticities with respect to food prices and total expenditure with t-value at median in parenthesis (* $p < 0.05$). We estimate these elasticities for the third quartile sample with the panel model. Refer to the following full name of each food category for row title with †; 1. Rice 2. Maize 3. Wheat and other cereals 4. Cassava 5. Roots, Tubers and, other starches 6. Sugar 7. Pulses 8. Nuts and seeds 9. Vegetables 10. Fruit 11. Red meat 12. Poultry 13. Eggs 14. Fish and seafood 15. Dairy 16. Fats and oils 17. Coffee, tea, and cocoa 18. Soft drink and juice 19. Other food Expenditure

Table C.4: Price and Total Expenditure Elasticities for the Fourth Quartile: Panel Model

	Food Category																			Expenditure
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. Rice	-1.37*	0.34	-0.09	0.42	-0.26	0.05	0.10	0.04	-0.04	0.24	0.22	0.09	0.12	-0.09	-0.06	-0.18*	0.03	-0.02	-0.02	0.88*
	(-4.61)	(1.35)	(-0.66)	(1.64)	(-1.28)	(1.17)	(1.03)	(0.45)	(-0.90)	(1.46)	(1.47)	(0.60)	(1.74)	(-0.78)	(-0.81)	(-2.08)	(1.16)	(-0.28)	(-1.37)	(4.05)
2. Maize	0.26	-0.91	0.23	-0.45	0.17	0.18*	0.03	-0.74	-0.07	0.05	0.10	-0.55	0.03	0.02	0.05	0.18	0.07	0.20	0.04*	0.82
	(1.00)	(-1.81)	(1.13)	(-1.26)	(0.76)	(1.99)	(0.45)	(-1.78)	(-1.11)	(0.28)	(0.82)	(-1.14)	(0.53)	(0.30)	(0.73)	(1.84)	(1.86)	(1.60)	(1.99)	(1.26)
3. Wheat†	-0.13	0.40	-1.32*	0.58*	0.40	-0.17*	0.12	0.59*	-0.03	0.01	-0.41*	0.59	0.02	0.29*	-0.28*	-0.07	0.04	-0.24	-0.00	1.04*
	(-0.77)	(1.51)	(-4.84)	(2.12)	(1.65)	(-2.00)	(0.92)	(2.36)	(-0.63)	(0.08)	(-2.03)	(1.90)	(0.26)	(2.00)	(-2.13)	(-0.94)	(1.39)	(-1.85)	(-0.31)	(4.98)
4. Cassava	0.70	-0.70	0.71*	-1.66*	-0.66	-0.05	-0.34	0.13	0.14	0.20	0.70	0.06	0.10	0.19	0.53	0.11	-0.08	-0.18	0.03	-0.09
	(1.94)	(-1.40)	(2.26)	(-3.62)	(-1.61)	(-0.82)	(-1.74)	(1.00)	(1.33)	(0.92)	(1.82)	(0.14)	(1.28)	(1.13)	(1.95)	(1.11)	(-1.90)	(-1.22)	(1.65)	(0.37)
5. Roots†	-0.34	0.18	0.31	-0.57	-2.06*	-0.06	-0.27	-0.03	0.18	-0.42	0.28	-0.32	-0.19	-0.42	0.33	0.20	-0.02	0.01	0.02	1.44*
	(-1.30)	(0.66)	(1.23)	(-1.65)	(-3.69)	(-1.02)	(-1.50)	(-0.46)	(1.70)	(-1.49)	(1.37)	(-0.96)	(-1.53)	(-1.85)	(1.65)	(1.76)	(-0.57)	(0.08)	(1.19)	(2.91)
6. Sugar	0.12	0.59*	-0.36*	-0.15	-0.10	-0.50*	0.23	0.02	0.15	-0.07	-0.08	-0.21	0.02	-0.05	-0.02	0.08	-0.09	0.05	-0.01	0.96*
	(1.23)	(3.51)	(-2.40)	(-1.29)	(-0.84)	(-2.05)	(1.43)	(0.28)	(1.77)	(-0.53)	(-0.70)	(-1.36)	(0.33)	(-0.53)	(-0.39)	(0.76)	(-1.78)	(0.77)	(-0.62)	(5.94)
7. Pulses	0.15	0.07	0.15	-0.44	-0.39	0.15	-1.27*	-0.22	0.14	0.03	-0.04	-0.27	0.09	0.40*	0.25	-0.17	0.05	-0.08	-0.09*	0.98*
	(0.82)	(0.42)	(0.93)	(-2.09)	(-1.72)	(1.32)	(-3.93)	(-1.22)	(1.29)	(0.22)	(-0.31)	(-1.16)	(0.72)	(2.11)	(1.79)	(-1.50)	(0.91)	(-0.72)	(-2.30)	(3.45)
8. Nuts†	0.13	-2.10*	1.01	0.19	-0.02	0.02	-0.26	-1.23*	-0.03	0.62	-0.70	0.20	-0.31	0.63	-0.53	0.01	0.01	-0.20	0.07	0.96*
	(0.43)	(-2.01)	(1.93)	(0.64)	(-0.00)	(0.25)	(-1.08)	(-2.95)	(-0.38)	(1.41)	(-1.66)	(0.44)	(-1.40)	(1.74)	(-1.73)	(0.07)	(0.12)	(-1.06)	(1.75)	(2.13)
9. Vegetables	-0.05	-0.11	-0.02	0.09	0.33*	0.10*	0.15	-0.02	-0.77*	-0.25*	0.21*	0.15	0.07	0.13*	0.06	-0.11*	-0.02	0.07	-0.03*	0.60*
	(-0.16)	(-0.88)	(-0.02)	(1.39)	(3.15)	(2.04)	(1.71)	(-0.00)	(-9.46)	(-2.30)	(2.45)	(1.37)	(1.47)	(2.68)	(1.01)	(-2.09)	(-0.61)	(1.34)	(-2.35)	(4.72)
10. Fruit	0.44	0.16	0.03	0.17	-0.52	-0.03	0.05	0.45	-0.23	-0.52	0.56	0.21	0.19	0.43	0.09	-0.04	-0.00	0.08	-0.04	0.55
	(1.23)	(0.53)	(0.17)	(0.60)	(-1.21)	(-0.31)	(0.32)	(1.31)	(-1.57)	(-1.62)	(1.55)	(0.71)	(1.24)	(1.51)	(0.72)	(-0.38)	(0.08)	(0.70)	(-1.51)	(1.50)
11. Red meat	0.16	0.07	-0.33*	0.39*	0.24	-0.04	-0.04	-0.30*	0.07	0.28	-1.43*	-0.21	0.24*	-0.11	-0.64*	-0.11*	0.10*	-0.27*	-0.04*	1.53*
	(0.79)	(0.46)	(-2.76)	(2.08)	(1.56)	(-0.95)	(-0.61)	(-2.42)	(1.68)	(1.84)	(-6.97)	(-1.53)	(2.99)	(-1.37)	(-4.06)	(-2.54)	(3.48)	(-2.90)	(-2.75)	(8.19)
12. Poultry	0.09	-0.62	0.40	0.01	-0.24	-0.08	-0.14	0.08	0.07	0.11	-0.18	-0.46*	-0.04	0.28	0.01	0.02	-0.05	0.00	0.02	1.28*
	(0.35)	(-1.49)	(1.64)	(-0.10)	(-0.83)	(-1.28)	(-1.18)	(0.31)	(0.90)	(0.48)	(-0.95)	(-2.13)	(-0.61)	(1.70)	(0.25)	(0.22)	(-1.50)	(0.05)	(1.57)	(3.17)
13. Eggs	0.53	0.11	0.07	0.29	-0.74	0.03	0.22	-0.62*	0.17	0.49	1.10*	-0.21	-0.71*	0.23	-0.27	-0.17	0.09	0.02	-0.09	1.43*
	(1.71)	(0.44)	(0.20)	(0.94)	(-1.70)	(0.28)	(0.76)	(-2.06)	(1.00)	(1.51)	(2.53)	(-0.73)	(-2.87)	(0.96)	(-1.28)	(-1.10)	(1.25)	(0.10)	(-1.93)	(4.36)
14. Fish†	-0.14	0.09	0.37*	0.16	-0.53*	-0.02	0.38	0.45*	0.13	0.42*	-0.12	0.51*	0.09	-1.15*	0.16	-0.19	-0.03	-0.07	-0.02	0.72*
	(-0.57)	(0.55)	(2.09)	(0.87)	(-2.05)	(-0.37)	(1.85)	(2.21)	(1.64)	(2.03)	(-0.94)	(2.01)	(1.25)	(-8.31)	(1.46)	(-1.93)	(-1.08)	(-0.84)	(-1.44)	(3.50)
15. Dairy	-0.13	0.06	-0.35*	0.53*	0.45*	-0.02	0.21*	-0.38*	0.02	0.06	-1.08*	0.01	-0.09	0.13	-1.26*	-0.05	0.00	-0.13	0.01	1.31*
	(-1.45)	(0.43)	(-3.12)	(2.67)	(2.43)	(-0.64)	(2.00)	(-2.87)	(0.52)	(0.16)	(-5.45)	(-0.06)	(-1.56)	(1.49)	(-5.78)	(-0.86)	(0.29)	(-1.69)	(0.59)	(8.65)
16. Fats†	-0.62	0.72	-0.18	0.17	0.60	0.09	-0.31	-0.01	-0.25	-0.10	-0.34	0.07	-0.12	-0.40	-0.08	-0.20	0.34*	-0.18	-0.04	1.03*
	(-1.73)	(1.82)	(-0.86)	(0.85)	(1.76)	(0.63)	(-1.27)	(-0.06)	(-1.64)	(-0.52)	(-1.45)	(0.23)	(-0.93)	(-1.66)	(-0.60)	(0.29)	(2.10)	(-1.33)	(-0.83)	(2.29)
17. Coffee†	0.18	0.70*	0.24	-0.50*	-0.09	-0.25	0.20	0.02	-0.10	-0.04	0.87*	-0.38	0.17	-0.17	0.03	0.83*	-1.17*	-0.00	0.04	1.08*
	(0.97)	(2.91)	(1.41)	(-2.52)	(-0.50)	(-1.85)	(0.91)	(0.08)	(-0.96)	(-0.14)	(3.30)	(-1.67)	(1.33)	(-1.29)	(0.38)	(3.64)	(-11.49)	(-0.03)	(0.93)	(5.17)
18. Soft drink†	-0.14	0.65	-0.60	-0.42	-0.00	0.03	-0.17	-0.29	0.06	0.11	-0.90*	-0.03	0.01	-0.19	-0.26	-0.19	-0.01	-1.06*	0.05	1.80*
	(-0.66)	(1.64)	(-1.88)	(-1.58)	(-0.02)	(0.36)	(-0.89)	(-1.29)	(0.48)	(0.28)	(-2.32)	(-0.57)	(-0.13)	(-1.21)	(-1.45)	(-1.81)	(-0.25)	(-3.70)	(1.74)	(3.68)
19. Other food	-0.33	0.66*	-0.04	0.30*	0.27	-0.06	-0.71*	0.49*	-0.31*	-0.40*	-0.52*	0.37*	-0.31*	-0.23	0.09	-0.17	0.09	0.28	-1.01*	0.77*
	(-1.44)	(3.52)	(-0.15)	(2.03)	(1.66)	(-0.57)	(-2.71)	(3.36)	(-2.32)	(-2.46)	(-2.51)	(2.19)	(-2.20)	(-1.67)	(0.78)	(-0.94)	(1.00)	(2.18)	(-8.82)	(3.74)

Notes: This table shows consumption quantity elasticities with respect to food prices and total expenditure with t-value at median in parenthesis (* $p < 0.05$). We estimate these elasticities for the fourth quartile sample with the panel model. Refer to the following full name of each food category for row title with †; 1. Rice 2. Maize 3. Wheat and other cereals 4. Cassava 5. Roots, Tubers and, other starches 6. Sugar 7. Pulses 8. Nuts and seeds 9. Vegetables 10. Fruit 11. Red meat 12. Poultry 13. Eggs 14. Fish and seafood 15. Dairy 16. Fats and oils 17. Coffee, tea, and cocoa 18. Soft drink and juice 19. Other food Expenditure